Satellite-Based Fog Detection

A Dynamic Retrieval Method for Europe Based on Machine Learning

Doctoral Thesis by Sebastian Egli

Satellite-Based Fog Detection: A Dynamic Retrieval Method for Europe Based on Machine Learning

kumulative Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem Fachbereich Geographie
der Philipps-Universität Marburg
vorgelegt von

Sebastian Felix Egli
aus Schaffhausen, Schweiz

Marburg / Lahn 2019

Vom Fachbereich Geographie
der Philipps-Universität Marburg (Hochschulkennziffer: 1180)
als Dissertation am 25.04.2019 angenommen.

Erstgutachter: Prof. Dr. Jörg Bendix
Zweitgutachter: Prof. Dr. Jan Cermak

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Christian Opp
Prof. Dr. Thomas Brenner
Prof. Dr. Bernhard Seeger

Tag der mündlichen Prüfung: 07.05.2019

Für Resie

Preface

Bringing this work to a successful end would not have been possible without the help of my colleagues, friends and family.
First of all, I would like to thank my supervisor Jörg Bendix who not only provided the financial and organizational framework for this project but also granted me the freedom to develop my ideas and who gave me orientation when I struggled with the many imponderables of fog research.
I would also like to thank Boris Thies for his tireless and ambitious support over the years. Be it during long days in the office or during even longer nights outside in the fog. I could not have wished for a better mentor.
I am also grateful to Sebastian Achilles for his technical support in the field and to Johannes Drönner for sharing his profound programming knowledge. With your help I managed to solve the numerous problems I was facing during data collection and algorithm development.
I also gratefully acknowledge the funding by the German Research Foundation under grants BE1780/40-1 as well as TH1531/4-1 and I want to thank the people at the German Weather Service and at the European Organisation for the Exploitation of Meteorological Satellites for providing the satellite and station data that form the basis of my work.
Because I know how important a positive working climate is, I would also like to thank the whole LCRS group for providing such a friendly and welcoming atmosphere. Many thanks for our fruitful (and sometimes heated) discussion rounds, long summer evenings at the Lahn, exuberant dinner and pub evenings, relaxed canoe trips as well as long hikes at the Edersee and through the Marburg countryside. I won't forget this rewarding time.
Finally, I want to say thanks to my friends, my family and my partner Resie. Thank you for giving me the supportive background that really made it possible for me to get to this point and thank you for being part of my life.

Sebastian Egli
Marburg, January 2019

Contents

List of Figures

List of Tables

List of Acronyms and Symbols

ABI	Advanced Baseline Imager
AVHRR	Advanced Very High Resolution Radiometer
CBA	Cloud Base Altitude [m a.g.l.]
CDP	Cloud Droplet Probe
CC	Cloud Cover
CL	Cloud Levels
CM-SAF	Satellite Application Facility on Climate Monitoring
CMA	Cloud Mask Product of CM-SAF
CP	Cloud Phase
CTH	Cloud Top Height [m a.g.l.]
CV	Cross Variogram
DEM	Digital Elevation Model
DFG	German Research Foundation
DWD	German Weather Service
DSD	Drop Size Distribution
ELV	Terrain Elevation [m a.m.s.l.]
EXCO	Extinction Coefficient [m^{-1}]
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAR	False Alarm Rate
FLS	Fog and Low Stratus
FMCW	Frequency Modulated Continuous Wave
GEO	Geosynchronous Equatorial Orbit
GOES	Geostationary Observational Environmental Satellites
GWS	General Weather Situation
HR	Hit Rate
HRV	High Resolution Visible band
HSS	Heidke Skill Score
HSS	Heidke Skill Score
INSAT-3D	Indian National Satellite - 3D
LEO	Low Earth Orbit
LIDAR	Light Detection an Ranging
LLO	Leave Location Out
LST	Low Stratus
LWC	Liquid Water Content [$g\,m^{-3}$]
LWP	Liquid Water Path [$g\,m^{-2}$]
MAD	Madogram
MAE	Mean Absolute Error
METAR	Meteorological Aviation Routine Weather Reports
MIR	Medium Infrared
ML	Machine Learning
MFG	Meteosat First Generation
MGD	Modified Gamma Distribution

MODIS	Moderate Resolution Imaging Spectroradiometer
MOR	Meteorological Optical Range [m]
MSG	Meteosat Second Generation
MTG	Meteosat Third Generation
NOAA	National Oceanic and Atmospheric Administration
NDSI	Normalized Difference Snow Index
N_t	Droplet Number Concentration [cm^{-3}]
NWP	Numerical Weather Prediction
OOB	Out Of Bag
PC	Percentage Correct [%]
PCV	Pseudo Cross Variogram
PDF	Probability Density Function
POD	Probability Of Detection
r_e	Effective Particle Radius [µm]
rh	Relative Humidity [%]
RF	Random Forest
RFE	Recursive Feature Elimination
ROD	Rodogram
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SOM	Self Organizing Map
ST	Smart Tether
STD	Standard Deviation
SVA	Satellite Viewing Angle [°]
SYNOP	Synoptic Weather Observations
T_B	Brightness Temperature [K]
TIR	Thermal Infrared
TPI	Topographic Position Index [m]
VAR	Variogram
VIS	Visibility [m]
WMO	World Meteorological Organization

Chapter 1

Introduction

1.1 Motivation

Fog has a strong effect on the environment in many aspects, e.g. as a water source for vegetation in otherwise arid regions like the Atacama desert in Chile (Lehnert et al., 2018) or the Namib desert in Southwest Africa (Andersen and Cermak, 2018; Henschel and Seely, 2008). It also plays a significant role as a major regulator of entire ecosystems, such as in the montane cloud forests of Taiwan (Li et al., 2015; Schulz et al., 2017), in the tropical lowland cloud forests of French Guiana (Obregon et al., 2014) or in the coast redwood region along the Pacific coast of California (Johnstone and Dawson, 2010).

Besides its many ecological impacts, it also has a direct influence on human life. For instance, fog often obstructs traffic on roads and airports due to sudden visibility reductions (Bendix et al., 2011). This leads to car accidents and capacity reduction or flight schedule delays at airports causing significant economic losses (Dietz et al., 2018; Kneringer et al., 2018; Robinson, 1989). Fog also destabilizes electricity grids by rapidly changing solar radiation conditions over photo-voltaic power plants (Köhler et al., 2017).

In addition to its influence potential on the economic domain, fog also acts as a modifier of the radiation balance of the atmosphere due to its high albedo. As a result, it prevents air layers close to the ground from warming up by reducing the absorption of incoming solar radiation, which leads to a strong restriction of vertical air circulation and thus to the persistence of inversion weather situations. The persistence of these inversions, in turn, can lead to severe health consequences when toxic substances like mercury (Weiss-Penzias et al., 2012) or sulfuric acids (Nemery et al., 2001) from the industry remain in the lower atmospheric layers in metropolitan areas for extended time spans. The most infamous of such morbidity-increasing fog events that were reported and scientifically investigated were the Meuse valley fog in Belgium in December 1930 and the London fog event of December 1952. Both fog events resulted in death tolls of several thousand people (Nemery et al., 2001).

Apart from these negative effects, fog can also be beneficial to humans, especially to those living in arid regions. Here, fog water can be accumulated using collector systems when other water supplies are exhausted (Schemenauer et al., 1988; Schemenauer and Cereceda, 1991).

The large number of fog influence factors shows that a comprehensive understanding of its causes and a precise mapping of the spatio-temporal distribution patterns are of great interest. Although the general fog-promoting atmospheric conditions are known, the microphysical properties of fog, especially in a differentiated view of its life cycle stages, are not conclusively understood. In order to close the data and knowledge gaps concerning fog microphysical properties and its meteorological drivers, an

in situ investigation during different fog life cycle stages is necessary. Although several surveys on meteorological and microphysical conditions in individual fog events have been conducted (Degefie et al., 2015; Fuzzi et al., 1992, 1998; Lu et al., 2010; Niu et al., 2010; Okita, 1962; Pinnick et al., 1978; Stolaki et al., 2015), most of these studies did only measure at or close to the ground level and therefore do not provide information on the complete fog column. Most important, life cycle stage information has not yet been taken into account during these investigations. Further surveys are thus needed to be able to make more reliable statements about fog properties over the complete life cycle.

In addition, existing fog mappings and climatological studies are carried out solely on the basis of station measurements (e.g. Avotniece et al., 2015; García-García and Zarraluqui, 2008; Srivastava et al., 2017; Troxler and Wanner, 1991; Vautard et al., 2009; Veljović and Vujović, 2018) because area-wide and spatially explicit data, required for a comprehensive spatial analysis, are not available. Since fog is a spatio-temporally very heterogeneous atmospheric phenomenon, investigations based on station data alone are insufficient to capture fog distributions representatively. A spatially explicit data set on fog distribution with a high temporal coverage is needed to overcome these shortcomings and to allow for a detailed examination of its occurrence patterns. Although there are several approaches for the generation of spatially explicit data sets on fog distribution based on satellite observations, these attempts are not yet able to differentiate between true ground-touching fog and lifted cloud layers at night and they were only developed for fogs resulting from radiative cooling. The main objective of this study therefore is the development of a satellite-based algorithm for a reliable derivation of fog occurrence independent of fog type and solar position.

This thesis thereby focuses on Europe as a study domain. Here, a special interest in a more detailed investigation of fog exists due to its large metropolitan agglomerations and its complex infrastructural networks regularly suffering from fog occurrence. Europe also has a well-maintained network of station observations as well as good satellite data coverage and thus provides a solid data basis for the generation and validation of newly developed fog products. Since many different fog types can occur in Europe, a model developed for this domain should also be basically extendable to other regions.

1.2 Object of Investigation

For a meaningful analysis a precise definition of the object of investigation is needed. Fog is internationally defined as an accumulation of suspended water droplets or ice crystals near the earth's surface, causing visibility to drop below 1 km (Glickman, 2000). It is distinguished from mist which causes visibility reductions between 1 km and 2 km and from haze which consists of dry particles or non-activated very small water droplets resulting in visibilities between 2 km and 5 km (Vautard et al., 2009).

This fog definition includes all fog types equally – regardless of the type of genesis, spatial extent and duration or chemical and microphysical composition. Depending on the field of application, different classifications are therefore used to do justice to the complex dynamics of different fog types. The most common classifications are based on differences in the formation process, the surrounding geographical characteristics and the physical state (Bruijnzeel et al., 2005; Cotton et al., 2011; Pérez-Díaz et al., 2017).

The distinction of fog events according to their physical state results in a subdivision into liquid water and solid ice fogs. However, since the latter only occurs at

temperatures below −30 °C (Gultepe et al., 2014), it is very rarely observed in Europe and therefore not considered any further in this study. Frequently mentioned fog types with regard to the surrounding geographical properties are sea- and coastal fog, mountain-, slope- and valley fog as well as urban fog (Bruijnzeel et al., 2005). Although the surrounding geography allows assumptions to be made about the fog formation processes to some extent, this nomenclature is of limited use when trying to describe the meteorological and microphysical properties of the respective fogs as well as the true underlying formation processes. Therefore, a distinction of fog types on the basis of their formation processes is more suitable in the context of this work. In general, fogs can be divided into three main formation categories which correspond to the following condensation mechanisms:

1. Air temperature decreases until the dew point is reached.
2. Humidity increases by addition of water vapor until the air is saturated.
3. Near-saturated air layers with different temperatures are mixed.

All three processes lead to condensation and thus – when occurring close to the ground – to fog formation when the dew point curve according to the Clausius-Clapeyron relation is exceeded (Callen, 1985). These formation processes are often not clearly distinguishable from each other and individual fog events may be influenced by all these processes simultaneously. However, in general, the first mechanism explains the formation of radiation and adiabatic cooling fogs, the second describes the formation of frontal fogs and the third describes the formation of advection fogs (Cotton et al., 2011).

Radiation fog usually forms at night during the winter season when atmospheric conditions are calm and stable. Radiative cooling of the earth's surface leads to a decrease of the air temperature in the lowermost air layers of the boundary layer. Due to the resulting temperature inversion, vertical air exchange is limited which prevents the cooling process from being weakened by an exchange with warmer air masses at the surface. When the air is cooled beyond the dew point, water starts to form small droplets around condensation nuclei. With a higher water vapor partial pressure, this process sets in sooner. This is why radiation fogs can most frequently be observed in river valleys like the upper Rhine or the Po valley where humidity values are generally higher. After sunrise and a resulting temperature increase of the lower air layers, radiation fog usually dissipates again quickly (Pérez-Díaz et al., 2017).

The term radiation fog is most often used in connection with ground or surface fog which can be observed in the lowermost parts of larger plains and basins. However, this category also includes raised or high inversion fog which is lifted from the lowermost terrain and thus is only perceived as fog in elevated areas, usually at the slopes of low mountain ranges adjacent to large depressions (Eugster, 2008; Scherrer and Appenzeller, 2014). The term valley fog also belongs to the radiation fog category. It is used for fog that forms in the valleys of mountainous areas in which its formation is often additionally promoted by katabatic flows (Hang et al., 2016).

Another fog type that forms due to a general cooling process of the air masses is fog that forms as a result of adiabatic cooling due to orographically induced uplift of advected air masses. It is generally referred to as upslope or orographic fog (Bruijnzeel et al., 2005) and can usually be observed at the slopes along the windward side of higher mountain ranges when adiabatic cooling results in saturation and subsequently in condensation. Orographic fog must not be confused with mountain fog, which is caused by the transport of clouds into mountainous regions causing low visibilities in

exposed parts of the terrain, e.g. mountain peaks (Eugster, 2008). This type of fog is a special type in so far as the condensation process has already taken place at another location and the air masses are only relocated afterwards.

Fog resulting from an addition of moisture is generally not coupled to terrain properties but to the prevailing large scale weather situation. It occurs during the passage of warm (prefrontal fog) and cold fronts (postfrontal fog) when precipitation from upper air layers falling into underlying subsaturated air masses causes the water vapor saturation pressure to be exceeded in these layers (Cotton et al., 2011).

Advection fog forms when warm and near-saturated air masses are transported over colder surfaces. In contrast to radiation fog, it is therefore dependent on the presence of steady winds. The temperature of the advected air masses decreases when transported over the cold surface which causes condensation to set in as soon as the dew point is exceeded. Sea and coastal fogs are the most common types of this category. They usually develop when cold and warm ocean currents are in close proximity and wind fields are directed from warm to cold surface regions (Bruijnzeel et al., 2005). An overview of all listed fog types is given in Figure 5.2.

In addition to the meteorological formation mechanisms, the various fog types also differ with respect to their microphysical properties. Radiation fogs are characterized by small droplets with drop radii generally not exceeding 10 µm. Due to the small drop size, liquid water content (LWC) is also relatively small with values far below $1\,\mathrm{g\,m^{-3}}$ (Liu et al., 2012; Okita, 1962; Pinnick et al., 1978) and the vertical distribution of LWC in radiation fogs is believed to principally follow a subadiabatic profile (Cermak and Bendix, 2011).

Although the investigation of other fog types has received less attention so far, the existing studies show that their microphysical properties differ considerably from those of radiation fogs. Following Eugster (2008), advection fogs are characterized by larger drop sizes up to 50 µm and it is assumed that highest LWC values are reached in the middle of the vertical fog column of these fog types (Cotton et al., 2011). Frontal fogs, in turn, differ in their microphysical properties from all other fog types due to very small drop counts with values generally below $10\,\mathrm{cm^{-3}}$ (Liu et al., 2012).

1.3 Existing Fog Retrieval Methods

A lot of effort has been put into the development of fog retrieval and nowcasting techniques by means of satellite data exploitation in order to get spatial information on fog distribution. For a comprehensive overview see Gultepe et al. (2007) and Jacobs et al. (2008).

The potential of fog retrieval techniques based on low earth orbiting (LEO) satellite systems such as the AVHRR (Advanced Very High Resolution Radiometer) aboard the NOAA satellites and MODIS (Moderate Resolution Imaging Spectroradiometer) aboard the Terra & Aqua satellites has been extensively investigated (Bendix and Bachmann, 1991; Bendix, 2002; Bendix et al., 2005, 2006; Eyre et al., 1984; Turner et al., 1986). These methods are based on assumptions about fog microphysics, the properties of which are reflected in the characteristics of certain spectral band combinations. Of special relevance is the scattering and reflection behaviour of small droplets dominating fog and low stratus (FLS) in the medium (MIR) and thermal infrared range (TIR) in comparison to other cloud types with larger drop sizes. FLS shows strong reflectivities during the day in the MIR range around 3.9 µm compared to TIR range around 10.8 µm. At night, in contrast, FLS differs from land and other clouds by its low emissivity properties in the MIR range. Bendix (2002) makes use of

this fact and takes the brightness temperature difference between band 3 and 4 of the AVHRR instrument to differentiate FLS from other cloud types. A similar approach was used for the MODIS instrument where multiple spectral threshold-tests are performed on band 1-7. The threshold values for these tests were previously computed using a radiative transfer model based on input values of typical optical and geometrical FLS properties (Bendix et al., 2006). Schulz et al. (2016) employed a different approach which is not dependent on assumptions about the microphysical properties of fog. Instead, their technique uses negative correlations between terrain height and optical cloud thickness which results from cloud intercepting terrain. However, this approach is only applicable to adiabatic fog types and was only tested in Taiwan. Since the approach also requires heterogeneous terrain, it is not directly applicable to the complete domain of Europe.

Regardless of the exact methodology, all LEO-based approaches suffer from the disadvantage that they can only capture limited areas at a time and that the temporal resolution is poor due to low return rates. This means that it is not possible to capture large regions spatio-temporally representatively and that daily variations in fog occurrence cannot be observed accurately.

Geostationary satellite systems (GEOs), on the other hand, are more suitable to provide continuous information on fog distribution over long time periods and large areas with a high temporal resolution while assuring a frequent measurement interval. Due to the improvement of the spectral resolution of the new generations of GEO satellites, the listed methodologies have also been successfully transferred to the ABI system (Advanced Baseline Imager) aboard the US-operated GOES satellites (Geostationary Observational Environmental Satellites), the SEVIRI system (Spinning Enhanced Visible and Infrared Imager) aboard the European MSG (Meteosat Second Generation) satellites and the INSAT-3D Imager (Calvert and Pavolonis, 2010; Cermak and Bendix, 2007; Chaurasia and Jenamani, 2017; Ellrod, 1995).

The principle for FLS detection in these approaches is basically the same as for LEO satellites. Band combinations from the MIR and TIR are used to differentiate between cloud types according to cloud height and microphysical properties. Ellrod (1995) produces nightly band difference images of GOES 7/8 bands 3 and 4 and uses simple parameterizations to differentiate FLS from other cloud type areas. Cermak and Bendix (2007) and Chaurasia and Jenamani (2017) apply histogram-based approaches onto MSG-SEVIRI bands 4 and 9 and onto INSAT-3D bands 3 and 5 respectively. All mentioned bands of GOES, MSG and INSAT correspond to the same spectral wavelength regions between 3.9 µm and 10.8 µm. Using this technique, they are able to calculate thresholds used for the differentiation of FLS and other clouds dynamically for each scene.

However, all these methods share two common problems: Only radiative or advective fog types are accounted for and a differentiation between low stratus (LST) and true ground touching fog is not adequately dealt with. This differentiation still is the central problem in the field of satellite-based fog detection because satellites can only observe the atmosphere from above. Information on the cloud or fog base can therefore only be inferred indirectly from the signals arriving at the satellite sensor. Several methods for the derivation of geometric cloud thickness as a prerequisite for the distinction between LST and ground touching fog have been proposed using pseudo-sounding approaches (Chang, 2003) or linear and adiabatic assumptions regarding the liquid water profile (Brenguier et al., 2000; Hutchison, 2002; Iwabuchi and Hayasaka, 2002). However, these methods do not perform well for FLS due to its unique microphysical properties (Bendix et al., 2005; Cermak et al., 2006).

Cermak and Bendix (2011) have recently developed a technique based on MSG-SEVIRI data that takes greater account of the microphysical features in fogs by applying a subadiabatic LWC profile in the fog derivation process. They use the method proposed by Kawamoto and Nakajima (2001) to estimate liquid water path (LWP) and cloud top heights (CTH) which are then used to calculate cloud base altitudes (CBA). These CBAs are subsequently merged with a digital elevation model (DEM) to derive areas of fog occurrence.

1.4 Aims and Hypotheses

In order to capture spatially explicit data on fog distribution with a high temporal resolution, the method developed by Cermak and Bendix (2011) is the only method that comes into closer consideration. Although their validation results show that the technique generally performs well, the scatter obtained during validation, however, points out that their microphysical assumptions might not hold for all fog life cycle stages. This assumption can also be confirmed by several studies in which it was shown that the droplet spectrum changes depending on the fog life cycle stage (e.g. Tampieri and Tomasi, 1976). Because of these uncertainties, it is necessary to review the microphysical assumptions made in the approach of Cermak and Bendix (2011). The first aim of this work therefore is to assess the microphysical and meteorological properties during different fog life cycle stages by means of in situ measurements over the entire fog column. On the basis of this investigation, a fog retrieval method can subsequently be developed that does justice to the complex dynamics of fog properties during different evolutionary stages.

Besides uncertainties in the microphysical approximations, several further restrictions prevent the scheme of Cermak and Bendix (2011) from being continuously applied to a large data set. First, the scheme was only developed for radiation fog conditions and the application to other fog types is not possible as other fog types show different illumination patterns resulting from their specific drop size spectra and water content profiles (Bruijnzeel et al., 2005). As a result, the climatologies created with this scheme do not take into account other fog types besides radiation fog. Second, the scheme can only be applied during daytime when the sun elevation is above 10°. This prevents the application of the scheme over the entire diurnal cycle. Especially in winter, when sun elevations remain low even during the day, this causes fog to be detectable only during short periods. Third, the methodology does not work when multiple cloud layers obstruct the satellite sensor's view because then the assumptions about the relations between fog microphysical properties and satellite signal cannot be used to infer fog information from the satellite data. Since a preliminary study of a combination of station observations and the CM-SAF cloud type product (Karlsson et al., 2011) showed that this condition occurs relatively frequently, the applicability of the procedure is further limited.

In summary, the physics-based fog retrieval algorithms have difficulties to adequately consider the dynamical interactions of all influence factors that cause fog formation and dissipation. This requires the development of a new fog detection scheme which can cope with the complex fog dynamics and the different fog types. The scheme should also be applicable during day and night and be able to cope with multiple cloud layers. This leads us to the second aim of this work: The implementation of a fog retrieval method capable of overcoming the above-mentioned difficulties which existing techniques are not yet able to cope with satisfactorily.

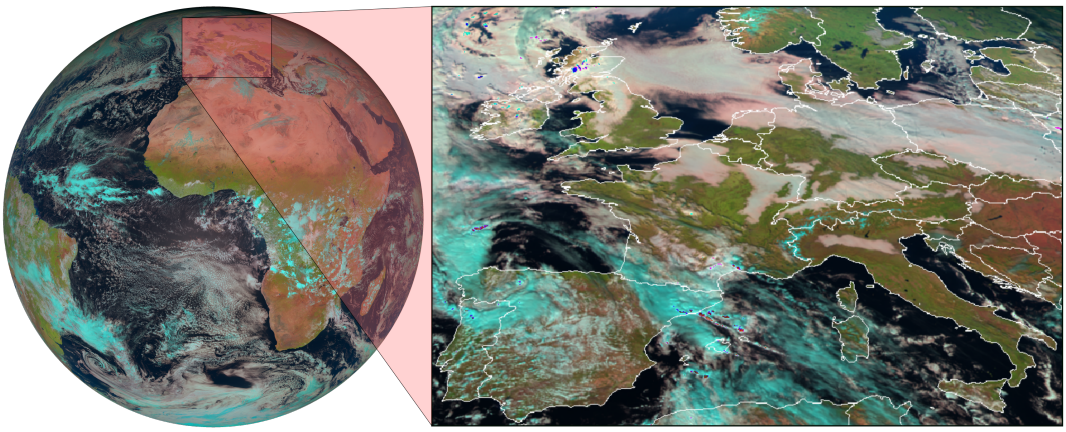

FIGURE 1.1: Natural color composite of November 15, 2011 - 12:00 UTC of the MSG full disk and the study domain of this thesis.

A growing number of studies are now focusing on the application of statistical machine learning (ML) techniques to derive atmospheric properties and climatological data sets from satellite data. Although various ML methods have already been applied in the context of rain retrieval and cloud segmentation (e.g. Drönner et al., 2018; Kühnlein et al., 2014), they have not yet been employed for the purpose of fog detection. However, ML methods should be well suited to handle the complex interactions of all fog influence variables and their manifestation in the satellite signal, since they make better use of the entire information content of the input variables provided. Therefore, the goal is to develop a dynamic ML based fog retrieval method that can be applied to geostationary satellite data operationally.

Since climatological assessments only make sense over longer time series, the fog retrieval approach should be developed for a system that on the one hand guarantees continuous spatio-temporal coverage and on the other hand reaches back into the past as far as possible. For Europe, the MSG-SEVIRI instrument is the ideal solution. An exemplary scene of November 15, 2011 at 12:00 UTC is depicted in Figure 1.1 showing a typical radiation fog situation over Europe. With a sub-satellite resolution of 3 km, it delivers continuous measurements every 15 minutes since 2004 with a suspected life time of the project reaching till 2024 when it will be replaced by Meteosat Third Generation (MTG). The data set created in this way has a spatial and temporal coverage that no other data set can compete with so far. It is therefore best suited as a basis for the creation of a fog climatology for Europe.

The resulting product from the application of this method can be used to make spatially explicit and temporally high-resolution analyses of fog dynamics for the first time. The third aim of this study consequently is to use the data obtained in this process for the creation of a baseline fog climatology. The data and information gained from this procedure should also be used to link fog distribution patterns with other atmospheric variables in order to analyze the large-scale control mechanisms of fog dynamics.

In summary, the following **research deficits** were identified:

- Existing fog retrievals are based on static subadiabatic assumptions about the vertical profile of fog microphysical properties. Due to a lack of in situ measurements over the entire fog column, it is not clear if these assumptions hold for all fog life cycle stages.
- Existing fog detection algorithms currently only work during the day and are not yet able to consider the complex meteorological and microphysical dynamics of fog sufficiently.
- There is no spatially explicit data set on fog distribution with a sufficiently high temporal resolution that allows for an analysis of the large-scale control mechanisms of fog.

In accordance with these research deficits, the **aims** of this thesis were formulated as follows:

- Assessment of microphysical and meteorological properties during different fog life cycle stages by means of in situ measurements over the entire fog column.
- Implementation of the first hybrid fog retrieval method capable of sufficiently deriving fog information for all fog types, regardless of the fog evolutionary stage and for the complete diurnal cycle.
- Creation of a baseline fog climatology and a subsequent analysis of the large-scale control mechanisms of fog dynamics with particular regard to the prevailing general weather conditions.

The following three **hypotheses** were developed in accordance with the research deficits and aims listed above:

H1 Microphysical properties change during fog evolutionary stages. Static subadiabatic approximations as currently used for the differentiation between fog and lifted cloud layers are therefore not suitable to cover the complete fog life cycle.

H2 To overcome the problems of existing fog retrievals, a dynamic hybrid approach is needed that combines satellite and station data via statistical machine learning methods in order to derive fog information regardless of fog type or evolutionary stage and for the complete diurnal cycle.

H3 On the basis of a fog climatological data set derived from geostationary satellite data using a machine learning technique, it is possible to analyze the large-scale control mechanisms of fog dynamics spatially explicitly over the complete diurnal and annual cycle, with special consideration of the prevailing general weather conditions.

1.5 Thesis Structure

In order to test these hypotheses, the following working packages were developed. An overview of the complete thesis structure is given in Figure 1.2.

WP1 In situ verification of microphysical assumptions in existing fog retrieval methods.

WP2 Dynamic adaptation of the FLS detection method as a precondition for a valid fog retrieval.

WP3 Development of a dynamic fog retrieval technique that uses a statistical machine learning approach.

WP4 Spatio-temporal fog pattern analysis and correlation with general weather situations.

WP1 was conceptualized to answer hypothesis **H1**. In an in situ measurement campaign at the Marburg ground truth and profiling station in Linden, Germany, meteorological properties and droplet spectra were recorded over the complete vertical fog profile during several radiation fog events (Chapter 2). To collect the data, a setup consisting of a combination of a ground-based radar unit and a visibility sensor as well as a tethered balloon borne system was used (Section 2.2.1). The results of this campaign were compared to theoretical assumptions used in current fog retrieval techniques in order to check for their applicability. An in-depth analysis showed that the theoretical assumption of a subadiabatic LWC profile in radiation fogs is valid in principle. However, strong variations during the different fog evolutionary stages make it necessary to adjust existing algorithms for a dynamic component that is able to account for changes in fog microphysics that manifest themselves in variations in the satellite signal (Section 2.3.3).

Thus, the main objective of this thesis was to develop a new fog retrieval technique based on satellite remote sensing data that would be able to cope with the complex meteorological and microphysical dynamics of fog and its spatio-temporal high variability in a satisfactory manner. Working package **WP2** and **WP3** were conceptualized to check the validity of hypothesis **H2** accordingly (Chapter 3 & 4).

The first step towards a valid fog retrieval is a correct detection of all cloud-covered pixels. Especially the accurate detection of FLS-covered areas that form due to radiative cooling is important, as this type is the most common cause of fog in the European lowlands. A correct detection of radiation FLS, however, is difficult, as signals arriving at the satellite strongly depend on the fog life cycle stage and the associated changes in the droplet spectrum, ground conditions and land cover, position of the sun and the satellite viewing angle.

For this reason, two processing steps from the method developed by Cermak and Bendix (2011) were extended by a dynamic component in working package **WP2**. The cloud detection method based on the 3.9 µm and 10.8 µm spectral ranges was adapted so that threshold values for cloud delineation were calculated separately for day/night and land/sea. In order to take changes in the sun zenith angle into account, these thresholds are recalculated for each scene and a moving-window method additionally ensures that changes in the satellite viewing angle can be taken into account in the histogram analysis underlying the threshold determination method. The 3.9 µm band partly overlaps with the lower CO2 absorption band centered around 4.2 µm. As CO_2 absorption varies depending on the length of the atmospheric column and therefore on the satellite viewing angle, in this thesis the small droplet proxy test which is used to

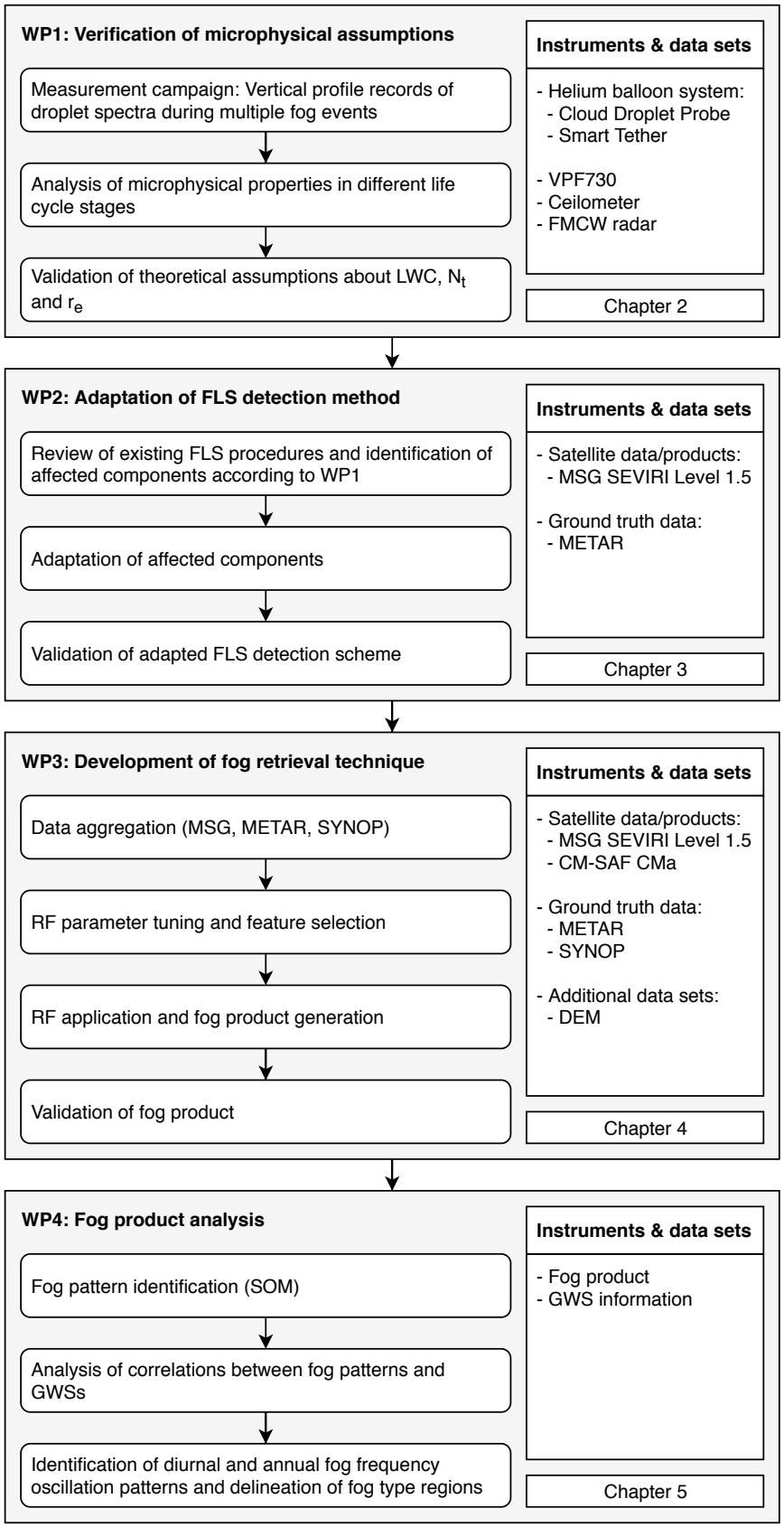

FIGURE 1.2: Thesis structure: Basic working packages of this study.

differentiate FLS from large droplet cloud types during daytime, is also extended by a moving window approach to ensure a fine grained threshold determination (Section 3.2.2).

The quality of the adapted FLS detection algorithm was checked in a detailed validation procedure with ground-based visibility and cloud cover data (Section 3.2.3). The results of this procedure were promising and the algorithm was deemed suitable for the derivation of a long term climatological data set (Section 3.3). A baseline FLS climatology was computed for Europe for the period between 2006 and 2015 and spatial distributions in FLS occurrence were analyzed. Seasonal and interannual variations in FLS distribution as well as long term trends were also investigated (Section 3.4).

Now that all cloud-covered areas including FLS can be determined with sufficient accuracy, the next challenge is to differentiate between ground-touching fog and lifted cloud patches. The physics-based fog retrieval technique of Cermak and Bendix (2011) computes fog microphysical properties from MSG bands in the visible and near infrared range and then uses this information to derive CBA values via iteratively fitting a theoretical LWC profile to LWP and CTH values. Unfortunately, the procedure is only applicable during daytime as information from the visible range is needed for the computation of the microphysical properties and the theoretical LWC profile used is only applicable to radiation fog. Furthermore, as mentioned above, the microphysical assumptions do not apply to all life cycle stages and if multiple cloud layers interfere with the satellite sensor's view, the algorithm cannot be used.

Working package **WP3** was therefore conceived to design a new algorithm that solves or bypasses the existing problems of fog detection regarding its microphysical dynamics and that is generally applicable to all fog types and to the complete diurnal cycle. Since ML approaches are particularly suitable for extracting the complex relationships of fog dynamics from the provided input data, a Random Forest (RF) approach was chosen (Breiman, 1996). RF models are computationally cheap ensemble learning methods. They run efficiently on large data sets, can easily handle vast numbers of input features and are relatively robust to noise and outliers. RFs also do not depend on the specification of an underlying data model and they can capture non-linear relations between input and target features. As RFs provide internal estimates of the different feature importances, they are also well suited to help filtering for relevant features concerning their prediction potential with respect to the target feature.

The algorithm was designed to predict CBA values for cloud covered regions and subsequently derive fog information via application of a statistically derived threshold separately for each pixel (Section 4.2.1). RFs, like any machine learning algorithm, need training data. The CBA information used for the training of the model was taken from METAR and SYNOP observations of 284 well distributed stations in Europe and was temporally interpolated onto the MSG scan times (Section 4.2.3 & 4.2.4). Additionally to the original MSG bands, band combinations, geostatistical texture features and terrain elevation measures were added to the input data set and a satellite viewing angle (SVA) correction was applied prior to the model training. In order to tweak the RF model towards better performance, a tuning data set was extracted from the global input data set which was then used for parameter tuning and feature selection. A separate model was then trained and applied to each MSG scene for the period from 2006 to 2015 (Section 4.2.5). The algorithm was validated using a leave-location-out (LLO) cross validation procedure. The results show that the new algorithm can outperform existing procedures, especially when considering its broad applicability to all fog types and to the complete diurnal cycle (Section 4.3).

In a last step, the computed fog product data set was used to analyze seasonal and diurnal fog frequency anomalies in Europe (Section 4.4).

After successfully developing a machine learning based fog retrieval technique and applying it to a large data set between 2006 and 2015, the third hypothesis can be tested. For this purpose, working package **WP4** was conceptualized. Because of its spatial coverage and the high temporal resolution, the fog product data set could be used to analyze fog distribution patterns in minute detail.

In order to gain an overview of the spatio-temporal fog distribution in Europe, representative fog patterns were retrieved using a self organizing map (SOM) approach. This procedure resulted in the clustering of all average daily fog frequency distributions into 30 distinct fog patterns (Section 5.3.1). Next, fog pattern occurrences were correlated to the occurrence of general weather situations (GWS) in order to reveal the large-scale meteorological control mechanisms of fog dynamics (Section 5.3.2). In a last step, diurnal and annual fog frequency oscillation patterns were identified and on the basis of the derived oscillation properties, regions with different fog types were delineated (Section 5.3.3).

In the following chapter, the field measurement campaign of **WP1** will be presented. Chapter 3 gives an overview of the dynamic adaptations of the low stratus detection of **WP2**. In Chapter 4 the new fog retrieval algorithm of **WP3** is explained in detail and in Chapter 5 the analysis of the generated fog product of **WP4** is presented.

Bibliography

Andersen, H. and Cermak, J. (2018). First fully-diurnal fog and low cloud satellite detection reveals life cycle in the Namib. *Atmospheric Measurement Techniques Discussions*, 20(July):1–15.

Avotniece, Z., Klavins, M., and Lizuma, L. (2015). Fog climatology in Latvia. *Theoretical and Applied Climatology*, 122(1-2):97–109.

Bendix, J. (2002). A satellite-based climatology of fog and low-level stratus in Germany and adjacent areas. *Atmospheric Research*, 64(1-4):3–18.

Bendix, J. and Bachmann, M. (1991). Ein operationell einsetzbares Verfahren zur Nebelerkennung auf der Basis von AVHRR-Daten der NOAA-Satelliten. *Meteorologische Rundschau*, 43:169–178.

Bendix, J., Eugster, W., and Klemm, O. (2011). Fog – boon or bane? *Erdkunde*, 65(3):229–232.

Bendix, J., Thies, B., Cermak, J., and Nauß, T. (2005). Ground Fog Detection from Space Based on MODIS Daytime Data—A Feasibility Study. *Weather and Forecasting*, 20(6):989–1005.

Bendix, J., Thies, B., Nauß, T., and Cermak, J. (2006). A feasibility study of daytime fog and low stratus detection with TERRA/AQUA-MODIS over land. *Meteorological Applications*, 13(2):111–125.

Breiman, L. (1996). *Out-Of-Bag Estimation*. PhD thesis, University of California.

Brenguier, J.-L., Pawlowska, H., Schüller, L., Preusker, R., Fischer, J., and Fouquart, Y. (2000). Radiative Properties of Boundary Layer Clouds: Droplet Effective Radius versus Number Concentration. *Journal of the Atmospheric Sciences*, 57(6):803–821.

Bruijnzeel, S., Eugster, W., and Burkard, R. (2005). Fog as a Hydrologic Input. In *Encyclopedia of Hydrological Sciences*, pages 559–582. John Wiley & Sons, Ltd, Chichester, UK.

Callen, H. B. (1985). *Thermodynamics and an Introduction to Thermostatistics*. John Wiley & Sons, 2 edition.

Calvert, C. and Pavolonis, M. (2010). GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document For Low Cloud and Fog. Technical report, NOAA NESDIS Center for Satellite Applications and Research.

Cermak, J. and Bendix, J. (2007). Dynamical nighttime fog/low stratus detection based on Meteosat SEVIRI data: A feasibility study. *Pure and Applied Geophysics*, 164(6-7):1179–1192.

Cermak, J. and Bendix, J. (2011). Detecting ground fog from space – a microphysics-based approach. *International Journal of Remote Sensing*, 32(12):3345–3371.

Cermak, J., Schneebeli, M., Nowak, D., Vuilleumier, L., and Bendix, J. (2006). Characterization of low clouds with satellite and ground-based remote sensing systems. *Meteorologische Zeitschrift*, 15(1):65–72.

Chang, F.-L. (2003). Retrieving vertical profiles of water-cloud droplet effective radius: Algorithm modification and preliminary application. *Journal of Geophysical Research*, 108(D24):4763.

Chaurasia, S. and Jenamani, R. K. (2017). Detection of Fog Using Temporally Consistent Algorithm With INSAT-3D Imager Data Over India. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(12):5307–5313.

Cotton, W. R., Bryan, G., and van den Heever, S. C. (2011). Fogs and Stratocumulus Clouds. In *International Geophysics*, volume 99, pages 179–242.

Degefie, D., El-Madany, T.-S., Hejkal, J., Held, M., Dupont, J.-C., Haeffelin, M., and Klemm, O. (2015). Microphysics and energy and water fluxes of various fog types at SIRTA, France. *Atmospheric Research*, 151:162–175.

Dietz, S. J., Kneringer, P., Mayr, G. J., and Zeileis, A. (2018). Forecasting Low-Visibility Procedure States with Tree-Based Statistical Methods. *Pure and Applied Geophysics*.

Drönner, J., Korfhage, N., Egli, S., Bendix, J., Freisleben, B., and Seeger, B. (2018). Fast Cloud Segmentation Using Convolutional Neural Networks. *Remote Sensing*, 10:1–8.

Ellrod, G. P. (1995). Advances in the Detection and Analysis of Fog at Night Using GOES Multispectral Infrared Imagery. *Weather and Forecasting*, 10(3):606–619.

Eugster, W. (2008). Fog Research. *Die Erde*, 139:1–10.

Eyre, J. R., Brownscombe, J. L., and Allam, R. J. (1984). Detection of fog at night using Advanced Very High Resolution Radiometer (AVHRR) imagery. *Meteorological Magazine*, 113(1346):266–271.

Fuzzi, S., Facchini, M. C., Orsi, G., Lind, J. A., Wobrock, W., Kessel, M., Maser, R., Jaeschke, W., Enderle, K. H., Arends, B. G., Berner, A., Solly, I., Kruisz, C., Reischl, G., Pahl, S., Kaminski, U., Winkler, P., Ogren, J. A., Noone, K. J., Hallberg, A., Fierlinger-oberlinninger, H., Puxbaum, H., Marzorati, A., Hansson, H.-C., Wiedensohler, A., Svenningsson, I. B., Martinsson, B. G., Schell, D., and Georgii, H. W. (1992). The Po Valley Fog Experiment 1989. *Tellus B*, 44(5):448–468.

Fuzzi, S., Laj, P., Ricci, L., Orsi, G., Heintzenberg, J., Wendisch, M., Yuskiewicz, B., Mertes, S., Orsini, D., Schwanz, M., Wiedensohler, A., Stratmann, F., Berg, O. H., Swietlicki, E., Frank, G., Martinsson, B. G., Günther, A., Dierssen, J. P., Schell, D., Jaeschke, W., Berner, A., Dusek, U., Galambos, Z., Kruisz, C., Mesfin, N. S., Wobrock, W., Arends, B., and Ten Brink, H. (1998). Overview of the Po valley fog experiment 1994 (CHEMDROP). *Contributions to atmospheric physics*, 71(1):3–19.

García-García, F. and Zarraluqui, V. (2008). A fog climatology for Mexico. *Erde*, 139(1-2):45–60.

Glickman, T. S. (2000). *Glossary of Meteorology (2nd ed.)*. American Meteorological Society, Boston, 2 edition.

Gultepe, I., Kuhn, T., Pavolonis, M., Calvert, C., Gurka, J., Heymsfield, A. J., Liu, P. S. K., Zhou, B., Ware, R., Ferrier, B., Milbrandt, J., and Bernstein, B. (2014). Ice fog in arctic during fram-ice fog project aviation and nowcasting applications. *Bulletin of the American Meteorological Society*, 95(2):211–226.

Gultepe, I., Tardif, R., Michaelides, S. C., Cermak, J., Bott, A., Bendix, J., Müller, M. D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., and Cober, S. G. (2007). Fog Research: A Review of Past Achievements and Future Perspectives. *Pure and Applied Geophysics*, 164(6-7):1121–1159.

Hang, C., Nadeau, D. F., Gultepe, I., Hoch, S. W., Román-Cascón, C., Pryor, K., Fernando, H. J. S., Creegan, E. D., Leo, L. S., Silver, Z., and Pardyjak, E. R. (2016). A Case Study of the Mechanisms Modulating the Evolution of Valley Fog. *Pure and Applied Geophysics*, 173(9):3011–3030.

Henschel, J. R. and Seely, M. K. (2008). Ecophysiology of atmospheric moisture in the Namib Desert. *Atmospheric Research*, 87(3-4):362–368.

Hutchison, K. D. (2002). The retrieval of cloud base heights from MODIS and three-dimensional cloud fields from NASA's EOS Aqua mission. *International Journal of Remote Sensing*, 23(24):5249–5265.

Iwabuchi, H. and Hayasaka, T. (2002). Effects of Cloud Horizontal Inhomogeneity on the Optical Thickness Retrieved from Moderate-Resolution Satellite Data. *Journal of the Atmospheric Sciences*, 59(14):2227–2242.

Jacobs, W., Nietosvaara, V., Bott, A., Bendix, J., Cermak, J., Michaelides, S., and Gultepe, I. (2008). *Short range forecasting methods of fog , visibility and low clouds - COST 722*. Office for Official Publications of the European Communities.

Johnstone, J. a. and Dawson, T. E. (2010). Climatic context and ecological implications of summer fog decline in the coast redwood region. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10):4533–8.

Karlsson, K. G., Meirink, J. F., and Stengel, M. (2011). Product User Manual CLOUDS. Technical report, The Satellite Application Facility on Climate Monitoring (CM SAF).

Kawamoto, K. and Nakajima, T. (2001). A global determination of cloud microphysics with AVHRR remote sensing. *Journal of Climate*, 14(9):2054–2068.

Kneringer, P., Dietz, S. J., Mayr, G. J., and Zeileis, A. (2018). Probabilistic Nowcasting of Low-Visibility Procedure States at Vienna International Airport During Cold Season. *Pure and Applied Geophysics*, (September).

Köhler, C., Steiner, A., Saint-Drenan, Y.-M., Ernst, D., Bergmann-Dick, A., Zirkelbach, M., Ben Bouallègue, Z., Metzinger, I., and Ritter, B. (2017). Critical weather situations for renewable energies – Part B: Low stratus risk for solar power. *Renewable Energy*, 101:794–803.

Kühnlein, M., Appelhans, T., Thies, B., and Nauß, T. (2014). Precipitation Estimates from MSG SEVIRI Daytime, Nighttime, and Twilight Data with Random Forests. *Journal of Applied Meteorology and Climatology*, 53(11):2457–2480.

Lehnert, L. W., Thies, B., Trachte, K., Achilles, S., Osses, P., Baumann, K., Bendix, J., Schmidt, J., Samolov, E., Jung, P., Leinweber, P., Karsten, U., and Büdel, B. (2018). A Case Study on Fog/Low Stratus Occurrence at Las Lomitas, Atacama Desert (Chile) as a Water Source for Biological Soil Crusts. *Aerosol and Air Quality Research*, 18(1):254–269.

Li, C.-F., Zelený, D., Chytrý, M., Chen, M.-Y., Chen, T.-Y., Chiou, C.-R., Hsia, Y.-J., Liu, H.-Y., Yang, S.-Z., Yeh, C.-L., Wang, J.-C., Yu, C.-F., Lai, Y.-J., Guo, K., and Hsieh, C.-F. (2015). Chamaecyparis montane cloud forest in Taiwan: ecology and vegetation classification. *Ecological Research*, 30(5):771–791.

Liu, D. Y., Niu, S. J., Yang, J., Zhao, L. J., Lü, J. J., and Lu, C. S. (2012). Summary of a 4-Year Fog Field Study in Northern Nanjing, Part 1: Fog Boundary Layer. *Pure and Applied Geophysics*, 169(5-6):809–819.

Lu, C., Niu, S., Tang, L., Lv, J., Zhao, L., and Zhu, B. (2010). Chemical composition of fog water in Nanjing area of China and its related fog microphysics. *Atmospheric Research*, 97(1-2):47–69.

Nemery, B., Hoet, P. H. M., and Nemmar, A. (2001). The Meuse Valley fog of 1930: An air pollution disaster. *Lancet*, 357(9257):704–708.

Niu, S., Lu, C., Liu, Y., Zhao, L., L??, J., and Yang, J. (2010). Analysis of the microphysical structure of heavy fog using a droplet spectrometer: A case study. *Advances in Atmospheric Sciences*, 27(6):1259–1275.

Obregon, A., Gehrig-Downie, C., Gradstein, S. R., and Bendix, J. (2014). The potential distribution of tropical lowland cloud forest as revealed by a novel MODIS-based fog/low stratus night-time detection scheme. *Remote Sensing of Environment*, 155:312–324.

Okita, T. (1962). Observations of the vertical structure of a stratus cloud and radiation fogs in relation to the mechanism of drizzle formation. *Tellus*, 14(3):310–322.

Pérez-Díaz, J., Ivanov, O., Peshev, Z., Álvarez-Valenzuela, M., Valiente-Blanco, I., Evgenieva, T., Dreischuh, T., Gueorguiev, O., Todorov, P., and Vaseashta, A. (2017). Fogs: Physical Basis, Characteristic Properties, and Impacts on the Environment and Human Health. *Water*, 9(10):807.

Pinnick, R. G., Hoihjelle, D. L., Fernandez, G., Stenmark, E. B., Lindberg, J. D., Hoidale, G. B., and Jennings, S. G. (1978). Vertical Structure in Atmospheric Fog and Haze and Its Effects on Visible and Infrared Extinction. *Journal of the Atmospheric Sciences*, 35(10):2020–2032.

Robinson, P. J. (1989). The Influence of Weather on Flight Operations at the Atlanta Hartsfield International Airport. *Weather and Forecasting*, 4(4):461–468.

Schemenauer, R. S. and Cereceda, P. (1991). Fog-water Collection in Arid Coasta Locations. *Ambio*, 20(7):303–308.

Schemenauer, R. S., Fuenzalida, H., and Cereceda, P. (1988). A Neglected Water Resource: The Camanchaca of South America. *Bulletin of the American Meteorological Society*, 69(2):138–147.

Scherrer, S. C. and Appenzeller, C. (2014). Fog and low stratus over the Swiss Plateau-a climatological study. *International Journal of Climatology*, 34(3):678–686.

Schulz, H. M., Li, C.-F., Thies, B., Chang, S.-C., and Bendix, J. (2017). Mapping the montane cloud forest of Taiwan using 12 year MODIS-derived ground fog frequency data. *PLOS ONE*, 12(2):17.

Schulz, H. M., Thies, B., Chang, S. C., and Bendix, J. (2016). Detection of ground fog in mountainous areas from MODIS (Collection 051) daytime data using a statistical approach. *Atmospheric Measurement Techniques*, 9(3):1135–1152.

Srivastava, S. K., Sharma, A. R., and Sachdeva, K. (2017). An observation-based climatology and forecasts of winter fog in Ghaziabad, India. *Weather*, 72(1):16–22.

Stolaki, S., Haeffelin, M., Lac, C., Dupont, J.-C., Elias, T., and Masson, V. (2015). Influence of aerosols on the life cycle of a radiation fog event. A numerical and observational study. *Atmospheric Research*, 151:146–161.

Tampieri, F. and Tomasi, C. (1976). Size distribution models of stratospheric particles in terms of the modified gamma function. *Archiv für Meteorologie, Geophysik und Bioklimatologie Serie A*, 25(1):47–54.

Troxler, F. and Wanner, H. (1991). Nebelkarten der Schweiz. *Geographica Helvetica*, 46(1):21–31.

Turner, J., Allam, R. J., and Maine, D. R. (1986). A case study of the detection of fog at night using channels 3 and 4 on the Advanced Very High Resolution Radiometer (AVHRR). *Meteorological Magazine*, 115:285–290.

Vautard, R., Yiou, P., and van Oldenborgh, G. J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nature Geoscience*, 2(2):115–119.

Veljović, K. and Vujović, D. (2018). Climatology of fog occurrence over a wide flat area in Serbia based on visibility observations. *International Journal of Climatology*, (September):1–14.

Weiss-Penzias, P. S., Ortiz, C., Acosta, R. P., Heim, W., Ryan, J. P., Fernandez, D., Collett, J. L., and Flegal, A. R. (2012). Total and monomethyl mercury in fog water from the central California coast. *Geophysical Research Letters*, 39(3):1–5.

Chapter 2

Vertical Distribution of Microphysical Properties in Radiation Fogs

This chapter has been published as: Egli, S., Maier, F., Bendix, J., & Thies, B. (2015). Vertical distribution of microphysical properties in radiation fogs - A case study. Atmospheric Research, 151, 130–145.

Received: 25 October 2013 / Accepted: 31 May 2014
https://doi.org/10.1016/j.atmosres.2014.05.027

Reprinted with permission from Elsevier.

Abstract

The present study investigates the validity of a theoretical liquid water content (LWC) profile in fog layers currently used for satellite based ground fog detection, with a special focus on the temporal dynamics during fog life cycle. For this purpose, LWC profiles recorded during two different fog events by means of a tethered balloon borne measurement system are presented and discussed. The results indicate a good agreement in trend and gradient between measured and theoretical LWC profiles during the mature stage of the fog life cycle. The profile obtained during the dissipation stage shows less accordance with the theoretical profile. To improve the agreement between theoretical and measured LWC profiles, the evolutionary stages during the fog life cycle should be incorporated. However, the variability within the prenoted measurements points out that more LWC profiles during a great variety of different fog events have to be collected for a well-justified adaptation of the theoretical LWC profile, considering fog life cycle phases in the future. In general, this underlines the existing knowledge gap regarding the vertical distribution of microphysical properties in natural fogs.

2.1 Introduction

Fog is internationally defined as a visible aggregate of small water droplets or ice crystals in the air, reducing horizontal visibility (VIS) to less than 1 km (Glickman, 2000). Its impact on the human sphere has increased significantly during the modern era - mainly due to obstructions of marine, air and road traffic or the solution and deposition of air pollutants (Bendix et al., 2011). As fog strongly influences the heat fluxes in the atmospheric boundary layer and the Earth's radiation budget by forming

persistent low level temperature inversions trapping air pollutants, it also often has a negative impact on the air quality in industrial agglomerations during inverted weather conditions (Nemery et al., 2001). All these effects ultimately lead to financial losses or even human casualties that are comparable to effects from other severe weather situations like tornadoes or winter storms (Gultepe et al., 2007). In contrast to its negative effects on the environment, fog is often considered as a positive element in ecology and hydrology as it can supply otherwise arid ecosystems with moisture (Bendix et al., 2011; Bruijnzeel et al., 2005; Pinto et al., 2001).

Hence, much effort has been put into the development of fore- and nowcasting methods by means of numerical models and satellite data (see Gultepe et al. (2007) and Jacobs et al. (2008) for an overview). The fog-forecasting-quality of numerical weather prediction (NWP) models is limited due to their low horizontal and vertical resolution as well as deficiencies in the parameterization of the microphysics (Bergot et al., 2007; Ronda et al., 2011). Several studies suggest that a better understanding of fog microphysics is needed to develop more accurate forecasting models (Gultepe et al., 2007; Pagowski et al., 2004; Tardif, 2007). The results of Terradellas and Bergot (2007) indicate a high sensibility of 1-D models to the incorporated microphysics and the vertical resolution. Since these models are computationally very intensive, parameterized versions of the detailed 1-D fog microphysics models can be used in 3-D models to improve the forecast results (Gultepe et al., 2006, 2007; Pagowski et al., 2004). As a consequence, several models explicitly simulate fog microphysics as e.g. PFOG (Bott and Trautmann, 2002). In several experiments, Thoma et al. (2012) could show that an initialization or nudging with observational data (visibility, vertical profiles of temperature and specific humidity) could significantly improve forecast quality. However, the authors stated that especially the integration of measured fog droplet spectra over the fog layer would be necessary to properly simulate fog formation, dissipation and fog vertical extent.

The potential of fog detection techniques based on polar orbiting satellite systems like NOAA/AVHRR (Advanced Very High Resolution Radiometer) and Terra & Aqua/MODIS (Moderate Resolution Imaging Spectroradiometer) has been widely investigated (Bendix, 2002; Bendix et al., 2005, 2006; Eyre et al., 1984; Turner et al., 1986) Due to improvements in the spectral resolution of the new generations of geostationary satellites, this methodology has also been successfully transferred to the GOES (Geostationary Observational Environmental Satellites) and the SEVIRI systems (Spinning Enhanced Visible and Infrared Imager) aboard MSG (Meteosat Second Generation) satellites (Cermak and Bendix, 2007; Ellrod, 1995). However, the common problem of these techniques persists: The distinction between low stratus (LST) layers and ground fog which is based on information about cloud top altitude and geometrical thickness. Various methods for the deduction of cloud geometrical thickness as a precondition to discriminate between LST and ground fog have been proposed using simple parameterizations (Ellrod, 1995), adiabatic approximations (Iwabuchi and Hayasaka, 2002) and pseudosounding approaches (Chang and Heston, 2002). However, these methods perform not always well for low stratus layers and fog (Bendix et al., 2005; Cermak, 2006). Hutchison (2002) estimated cloud thickness and cloud-base heights assuming constant vertical liquid water content (LWC) profiles for stratus clouds. Brenguier et al. (2000) applied an adiabatically stratified LWC profile to estimate cloud thickness. Minnis et al. (1997) developed an empirical relationship between cloud geometrical thickness and cloud optical thickness. Cermak and Bendix (2011) recently developed a technique on the basis of the European geostationary satellite system MSG-SEVIRI that takes greater account of the microphysical properties in fogs. In their model they assume a three-layered stratification of an LST/fog layer

with a shift from adiabatic to sub-adiabatic conditions in the lowermost part, sub-adiabatic conditions in the centre and a linear decrease of LWC to zero at the top of the layer, starting at about 80 % to 90 % of the fog vertical extent.

The validation results of Cermak and Bendix (2011) show an overall good performance of the ground fog detection methodology which points to the general validity of the underlying theoretical LWC profile. However, the scatter obtained during validation also pointed out that the assumptions might not hold for all fog events and/or fog life cycle phases. This is why the development of fog is often described as a typical life cycle with varying microphysical properties during consecutive evolutionary stages (Gultepe et al., 2007; Liu et al., 2011; Pilié et al., 1972, 1975b,a; Wendisch et al., 1998). Maier et al. (2013), for instance, classified three development stages with characteristic changes in the LWC at the ground level. However, it is obvious that LWC not only varies at the ground but also in the vertical profile during the different development stages. Thus, the validity of the remote sensing approach, but also of advanced simulation models like PAFOG might be strongly depending on the availability of varying vertical profiles of LST/fog properties.

Unfortunately, there is only little data concerning the vertical distribution of fog microphysics, particularly during the different life cycle stages, that would allow to investigate the validity of the theoretical LWC profile and/or properly initialize numerical models. In situ airborne measurements for the investigation of low level stratocumulus clouds (Hayasaka et al., 1995; Slingo et al., 1982; Wang et al., 2009) are hardly permitted and not even possible during fog situations. Studies using balloon-borne systems with suitable sensors for profile measurements are rare. Okita (1962), for instance, investigated the LWC profile of four radiation fogs in Japan. Pinnick et al. (1978) measured the vertical profile during a fog event in West Germany. However, the recorded profiles did not completely cover the whole fog column. During the Po Valley and the Chemdrop experiment the vertical structure during fog events was measured with a 50 m high tower (Fuzzi et al., 1992, 1998), also not capturing the whole vertical LST/fog profile. Beside the small data set about fog LWC profiles there is no information available about fog evolutionary stages during the different profile measurements. This information, however, is necessary to be able to account for the expected changes in the LWC profiles during the different development stages of the fog (Maier et al., 2013) and to verify the validity of the theoretical assumptions made by Cermak and Bendix (2011) especially throughout the whole fog life cycle.

The aim of the present study is to compensate the lack of data and knowledge by measuring LWC profiles during selected fog events with a special focus on their temporal evolution during the fog life cycle, particularly to assess the validity and applicability of the theoretical profile used in Cermak and Bendix (2011).

The paper is structured as follows: Section 2.2 describes the balloon borne measurement system together with other ground based measurement devices and introduces the necessary data processing steps. In section 2.3 first measurement results are presented and compared to the theoretical profile.

2.2 Instrumentation and Data Processing

2.2.1 Instrumentation

Vertical profiles of fog characteristics were recorded at the Marburg ground truth and profiling station in Linden, Germany (50.533°N, 8.685°E, 172 m a.m.s.l). The site is located in a rural valley, surrounded by flat hills extending up to 250 m a.m.s.l. These topographic conditions cause a high frequency of fog events during spring and

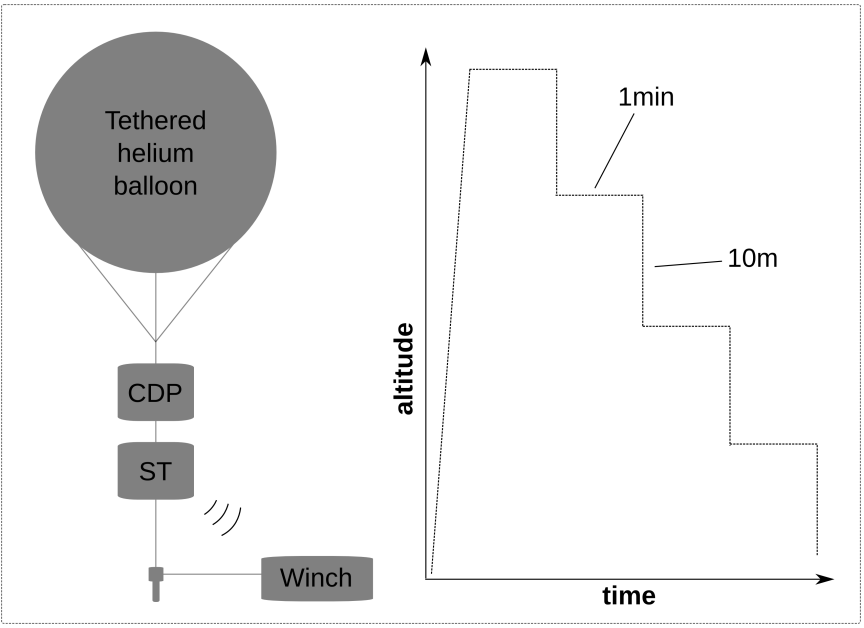

FIGURE 2.1: Schematic illustration of the methodological setup. CDP = Cloud Droplet Probe, ST = Smart Tether. Profiles were only measured during descension for 1 min on each 10 m level.

autumn months when weather situations with boundary-layer inversions frequently occur (Bendix, 2002; Schulze-Neuhoff, 1976). The measurements are based on a tethered balloon borne composition of meteorological and microphysical instruments, including a novel optical particle counter.

Via controlled ascending to the fog top (derived with the help of an FMCW radar device, described in detail below) and incremental descending to the ground level, meteorological and microphysical data profiles were recorded during the respective ground fog events with the instrumentation listed in Table 2.1. Ascent and descent of the measurement system were controlled with an electronic rope winch. A schematic illustration of the setup and the measurement procedure is given in Figure 2.1. Profiles were only recorded up to 250 m. This is due to safety reasons mainly caused by higher wind speeds above which could have damaged the instrumentation and possibly loosen the whole balloon system from the rope. To avoid a possible bias in the measurements due to drop expulsion in the lee of the balloon, only the data collected during descension was used. The instruments were held for 1 min on each altitude level and were iteratively lowered in 10 m steps. These intervals were suitable for an adequate representation of the DSD on each measurement altitude while also guaranteeing a sufficient vertical resolution of the profile. All operated instruments are described in detail in the following paragraphs.

Cloud Droplet Probe

Drop size distributions (DSD) were measured using a modified Cloud Droplet Probe (CDP) developed and manufactured by Droplet Measurement Technologies, Inc., Boulder, CO, USA. The instrument uses a 658 nm laser to illuminate particles in a specific volume of air and measures their size by capturing the intensity of the scattered light. A T2-03 Micro Diaphragm Pump, produced by the Parker Hannifin Corporation, Cleveland, OH, USA, was added to the aperture in order to provide

constant flow rates (PHC, 2010). In this way, the device was able to actively determine a measurement volume of 500 $cm^3\,min^{-1}$ which allowed for a direct derivation of absolute drop counts from the CDP measurements, thus making it suitable for a balloon borne application. Using this technique it was possible to detect drops at 30 intervals within the size range of 2 µm to 50 µm at a sampling frequency of 1 Hz. The manufaturer expects uncertainties in both, drop size and drop count measurements, not to exceed 20 % for a measured spectrum with decreasing errors towards smaller size bins (DMT, 2009, 2010). Also, records of similar forward scattering instruments like the FM-100 have been shown to suffice a derivation of LWC and N_t (Spiegel et al., 2012). To further account for measurement uncertainties, CDP records were averaged over one-minute intervals.

For a correct calibration of the modified CDP, prior to each studied fog event, the device's sizing accuracy was checked using glass beads with a known diameter. Additionally, the flowrate of the integrated pump system was adjusted using a flowmeter. In this way, both, average drop size and drop count measurements were ensured to be gathered under consistent conditions. To analyze the vertical fog structure, liquid water content (LWC), droplet number concentration (N_t) and effective particle radius (r_e) were calculated from the recorded DSD at each measurement height.

Smart Tether

Temperature, pressure, wind speed and relative humidity as well as the altitude of the instruments were measured with the wireless Smart Tether system distributed by Anasphere Inc., Bozeman, Montana, USA. With accuracies of 0.5 hPa, 0.5 °C, 1.0 $m\,s^{-1}$ and 3 % in pressure, temperature, wind speed and relative humidity (Anasphere, 2012), the instruments data quality is suitable for the precision needed in this study. In addition, altitude values were controlled with a barometric altimeter of a Garmin Oregon 450t GPS device with deviations not exceeding 3 m. The information recorded by the Smart Tether were reported to the ground station in real time via radio signal. In this way, the altitude of the instruments could be controlled from the ground and changes in wind speed could be monitored. When wind speed values exceeded 4 $m\,s^{-1}$ the balloon was pulled back.

VPF-730

Besides the tethered balloon observations, several instruments at ground level supported the study. For the distinction between fog and non-fog phases, visibility data was obtained with the HSS VPF-730 Visibility and present weather sensor developed at the Bristol Industrial and Research Associates Limited Company, England. It uses forward scatter meter technology to determine the extinction coefficient (EXCO) of a specific volume of air and calculates the meteorological optical range (MOR) on the basis of the measured data, thus providing visibility-data. The instrument is installed 2 m above ground and measures horizontal visibility every 20 s with error estimates below 10 % for visibility values between 0 km and 16 km (Biral, 2010, 2012).

FMCW Radar

Additional profile information about fog depth and density could be acquired using a Frequency Modulated Continuous Wave cloud radar profiler (FMCW radar) which has been developed by the STFC Rutherford Appleton Laboratory, Great Britain. The system operates at a frequency of 94 GHz which makes it highly sensitive to cloud and fog droplets while the signal's attenuation in the relevant atmospheric range

Instrument	Parameter	Range	Unit	Interval	Accuracy	Manufacturer
CDP	DSD	2 - 50	µm	1 s	20 %	DMT, USA
ST	temp	-55 - 125	°C	1 s	0.5 °C	Anasphere, USA
	press	0 - 1100	mbar	1 s	0.5 mbar	
	rh	0 - 100	%	1 s	3 %	
	ws	0 - 59	$m\,s^{-1}$	1 s	$1\,m\,s^{-1}$	
	alt	no spec.	m	1 s	no spec.	
VPF-730	VIS	0.01 - 75	km	20 s	10 %	BIRAL, GB
FMCW radar	a_{ft}	30 - 2000	m	10 s	3.9 m res.	RAL, GB
Ceilometer	a_{cb}	0 - 7600	m	20 s	5 m res.	Vaisala, Finland

TABLE 2.1: Instrumentation. DMT: Droplet Measurement Technologies; RAL: Rutherford Appleton Laboratories; BIRAL: Bristol Industrial and Research Associates Limited Company, CDP: Cloud Droplet Probe, ST: Smart Tether, VPF-730: Visibility and Present Weather Sensor. temp: Temperature, press: Air pressure, rh: Relative humidity, ws: wind speed, alt: Altitude, VIS: Horizontal visibility, a_{ft}: fog top altitude, a_{cb}: cloud base altitude

is relatively low. Furthermore, the near field of the device is significantly smaller than those of comparable pulse radar systems which results in reliable measurements already at 30 m above the instrument (Huggard et al., 2008). Bennett et al. (2009) and Huggard et al. (2008) also show that due to its high vertical resolution the device is able to capture cloud structures very well, with the saturation point not beeing exceeded even for multiple cloud layers. The temporal resolution of the device was set to 10 s with a vertical resolution of 3.9 m and a total range of 2.0 km. By this means, fog top altitude information was available during the profile measurements in real time.

Ceilometer

In order to get information about the cloud base height during fog formation and dissipation, a CL31 Ceilometer produced by Vaisala, Helsinki, Finland, was operated at the station. The device employs a pulsed diode laser LIDAR (light detection and ranging) technology to record cloud base altitudes of up to three cloud layers at 5 m vertical resolution for the range from 0 m to 6700 m (Vaisala, 2012). The complete instrumentation is listed in Table 2.1.

2.2.2 Data Processing

In a first preprocessing step, all parameters were averaged over 1 minute intervals in order to get a consistent temporal resolution throughout the whole data set and to account for measurement uncertainties.

As mentioned above, the microphysical parameters LWC, N_t and r_e were calculated from the measured DSD. In order to include the whole possible size range of fog droplets, the modified gamma distribution (MGD) was applied as a probability density function (PDF) to the raw data. The MGD has been shown to adequately fit the DSD of a vast variety of fog types and phases (Tampieri and Tomasi, 1976). It has therefore been chosen as a suitable representation of the drop size spectra recorded during this study. By applying the MGD, a more realistic representation of the range from 0 µm to 2 µm (which is not recorded by the CDP itself) and a smoothing of the discrete raw drop size distribution recorded by the device towards a continuous value distribution could be achieved. In this way, the expected total spectrum was

accounted for during the derivation of the fog microphysics and outliers in single size intervals were smoothed out. The MGD is given by Deirmendjian (1969) as

$$f(r) = ar^{\alpha}\exp(-br^{\gamma}) \tag{2.1}$$

where a, b and γ are positive real numbers and α is a positive integer. a is called the normalization factor as it ensures that the integral of the MGD over all radii equals the total drop count. The shape of the MGD is determined by b and α which are the slope and shape parameter respectively. It has two zeros, one at $r = 0$ and one at $r = \infty$. The MGD has been fitted to the DSD of each measurement using a least squares method implemented by Garbow et al. (1980) following the Levenberg-Marquardt algorithm. By this means the MGD parameters a, α, b and γ were derived.

According to Gradshteyn and Ryzhik (2007) the moments of the MGD can be written as

$$m_i = \frac{a}{\gamma} b^{-\left(\frac{\alpha+1+i}{\gamma}\right)} \Gamma\left(\frac{\alpha+1+i}{\gamma}\right) \tag{2.2}$$

with the gamma function

$$\Gamma(x) = \int_0^{\infty} t^{x-1}\exp(-t)dt \tag{2.3}$$

The 1st moment of the DSD gives total drop count N_t (cm^{-3}) while effective radius r_e (µm) can be derived via the 2nd and 3rd moment (Danne, 1996; Maier et al., 2012):

$$N_t = \frac{a}{\gamma} b^{-\left(\frac{\alpha+1}{\gamma}\right)} \Gamma\left(\frac{\alpha+1}{\gamma}\right) \tag{2.4}$$

and

$$r_e = \frac{b^{-\left(\frac{\alpha+4}{\gamma}\right)} \Gamma\left(\frac{\alpha+4}{\gamma}\right)}{b^{-\left(\frac{\alpha+3}{\gamma}\right)} \Gamma\left(\frac{\alpha+3}{\gamma}\right)} \tag{2.5}$$

Assuming a spherical drop shape and only liquid water within the fog layer, LWC (gm^{-3}) values can be calculated using the third moment of the DSD (Maier et al., 2012):

$$LWC = \frac{4}{3}\pi\rho_{H_2O}\frac{a}{\gamma} b^{-\left(\frac{\alpha+4}{\gamma}\right)} \Gamma\left(\frac{\alpha+4}{\gamma}\right) * 10^{-6} \tag{2.6}$$

where ρ_{H_2O} is the approximate water density of 1 g cm^{-3} at the expected temperatures.

Using these equations, the desired microphysical parameters were derived via the MGD for every single one-minute-average spectrum that was recorded by the CDP.

To determine the fog top altitude for the balloon borne measurements (section 2.3.3), visibility values were calculated from the DSD data using the following formula modified after DMT (2009) and Biral (2012):

$$VIS_c = 3\left(\pi \sum_{i=1}^{n} Qc_i {r_i}^2\right)^{-1} \tag{2.7}$$

with n being the count of radius intervals. Q is the extinction efficiency after van de Hulst (2012), c_i is the number concentration of hydrometeors in size category i and r_i is the average radius of size category i. The fog top was assumed to match the altitude at which the total drop size distribution diminished to a point where VIS_c was exceeding the 1 km threshold.

To account for the fog's evolutionary stages during the analysis of the profile records, 2 m VIS data of the different fog events were separated into different development phases using the breakpoint analysis proposed by Maier et al. (2013). Fog start and end date were defined as the first/last point in time when VIS decreased below 1 km.

To assess the validity of the theoretical LWC in comparison with the measured LWC profiles, the LWC and temperature values measured at ground level as well as the fog top height derived from DSD-deduced VIS_c values were used to compute the theoretical LWC profiles following the method described by Cermak and Bendix (2011). The model assumes a three-layered stratification of an LST/fog layer with the following characteristics.

1. The lowermost part of the cloud/fog layer shows a linear shift from adiabatic to sub-adiabatic conditions with increasing height.

2. Liquid water content (LWC) in the central region of low stratus and fog layers rises with altitude (Pruppacher et al., 1998). The emerging LWC profiles generally follow a sub-adiabatic gradient. This is due to the fact that most stratiform cloud formations are characterized by very weak turbulence which, in turn, leads to a higher saturation point and consequently to LWC values that diverge from the adiabatic profile (Betts, 1982).

3. Near the fog/stratus top r_e and thus LWC values decrease linearly to zero. In very low stratus layers this starts at about 80 % to 90 % of the cloud height and can be attributed to the entrainment of dry air at the top of the fog/stratus layer (Boers and Mitchell, 1994; Roach et al., 1982; Wieprecht et al., 2005).

For a detailed description of the concept please refer to Cermak and Bendix (2011).

2.3 Results and Discussion

In the following subsections, measurement results for two fog events are presented and compared to the theoretical profile. First, the vertical fog structure derived during the fog event on October 28/29 2011 (three profiles) is described. This will be followed by the presentation of the results obtained during the fog event on November 19/20 2012 (six profiles). All times are specified in UTC.

2.3.1 Observations on October 28/29, 2011

As presented in DWD (2013b), the synoptic weather regime in Europe during October 28/29, 2011 was characterized by a zonal ridge across Central Europe, following the classification of Hess and Brezowsky (1977), translated by James (2007). Troughs from a depression over the North Atlantic were not reaching over Central Europe and the weather in Linden was mainly influenced by a strong high pressure system over Eastern Europe (see Figure 2.2). As a consequence, clouds were absent and wind conditions remained calm at the study site during the whole measurement period.

A relatively thin and short-lived fog layer formed on October 28, 2011 at 23:23. It lasted 6 h 28 min with its top not exceeding 100 m altitude. Sunrise was at 06:11. The fog dissipated shortly before, at 05:50.

Figure 2.3 gives an overview of the boundary layer conditions at Linden during October 28/29, 2011. Due to the near field of the radar device, fog top altitudes were only plotted for fog tops above 30 m height. Before 01:52 and after 05:12, radar

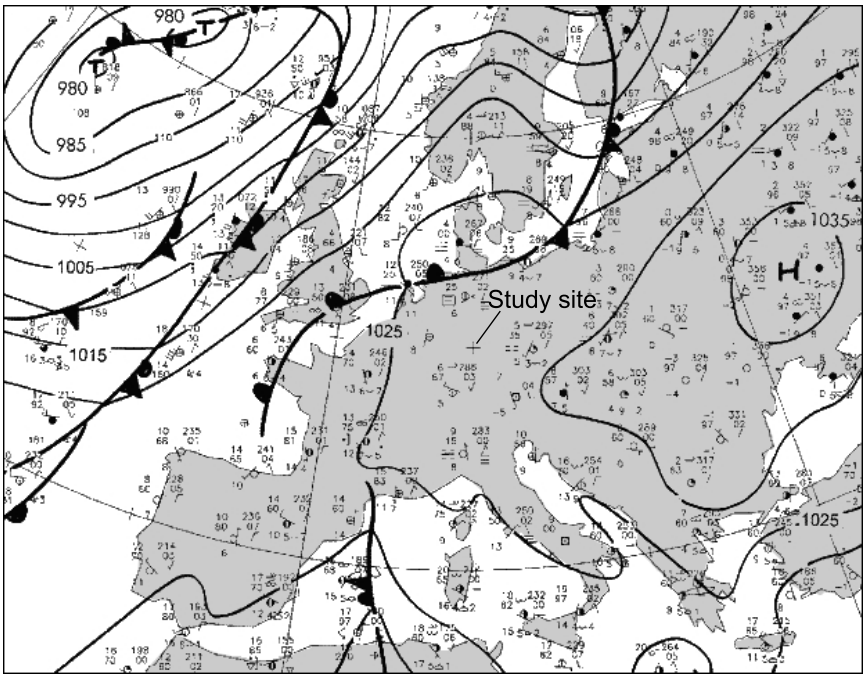

FIGURE 2.2: Synoptic weather situation in Central Europe for October 29, 2011: 00:00 UTC. Modified after DWD (2013a).

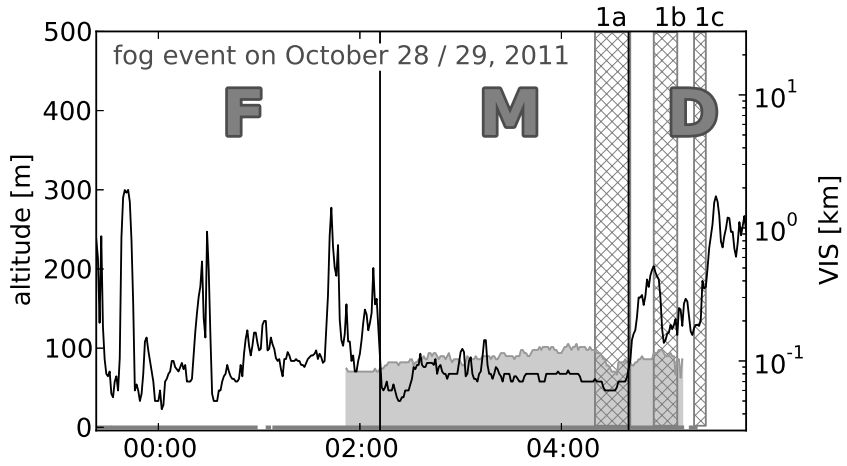

FIGURE 2.3: VIS (black line) & fog vertical extent (grey area) between 23:23 October 28, 2011 and 05:50 October 29, 2011. Fog top altitudes are derived from FMCW radar data (top grey line), cloud base altitudes from ceilometer data (bottom grey line). Hatched stripes mark periods of profile measurements and are labelled corresponding to the following profile plots. Measurement periods were 04:16-04:37 for 1a, 04:51-05:05 for 1b and 05:15-05:22 for 1c. Black vertical lines mark phase transitions. F: Formation, M: Mature fog, D: Dissipation.

reflectivity above 30 m did never exceed the noise floor of the device and it could be assumed that the fog layer did not outreach the near field altitude during these periods. Wherever the ceilometer was able to detect a reliable cloud base, it was recorded to be at ground level throughout the whole fog period. Following the terminology used in Maier et al. (2013), the first black vertical line marks the transition from formation to mature phase while the second line denotes the transition from mature phase to dissipation. While profile 1a was recorded at the end of the mature phase, dissipation has already started during the measurements of profile 1b and 1c.

The meteorological data along the fog vertical profiles presented in Figure 2.4 give first insights into the conditions of the lower atmosphere during this event. Temperature distributions indicate the existence of a strong ground-touching inversion layer up to approx. 110 m for profile 1a. In profile 1b the inversion has already been lifted from the ground and ranged from 50 m to 125 m altitude. Above the inversion layer, both profiles showed constant temperature values with averages of 7.7 °C in 1a and 8.3 °C in 1b. Until the recording of profile 1c, temperature values have further increased, only the uppermost part (above 150 m) showing a slight decrease. Isothermal conditions were thus only prevalent closely above the inversion layer at this point in time. The general temperature increase between the profile measurements can be ascribed to the increasing long-wave downward radiation reflected at the fog layer, especially during the mature stage when the fog layer was densest.

The rh profiles showed a steep decline towards higher altitudes within the inversion layer which can be attributed to the temperature rise at these heights. Above and below the inversion layer, the profiles showed comparably constant values with an exception in profile 1a where up to 60 m altitude rh gradually declined with height. A general increase of could be identified between profile 1a and 1b, likely resulting from an intensified evaporation of fog droplets due to the onset of the dissipation phase between both profile measurements. Profile 1c shows, that rh values generally declined afterwards which can be attributed to the contemporaneous temperature rise.

To give an overview of the drop size distributions of this fog, the average spectrum over all measured spectra of this event is given in Figure 2.5 together with its MGD fit. The derived MGD parameters are: $r_c = 0.99$, $\gamma = 0.32$, a = 1.24E13 and $\alpha = 8$. The graph shows large drop counts for the small size intervals and a steep decline towards larger drop sizes. This indicates that drop size distributions in this fog event were mainly governed by high counts of small particles.

The corresponding LWC profiles are depicted in Figure 2.6 (a). The absence of larger particles resulted in relatively low LWC values throughout the whole measurement period. Most values stayed below 0.1 g/m^3 which underlines the weakness of this fog event when compared to records of LWC in radiation fogs from other authors (Choularton et al., 1981; Okita, 1962; Pinnick et al., 1978; Wobrock et al., 1992) with values reaching up to 1.0 g/m^3. Standard errors of each one-minute sample are denoted as grey buffers. As LWC is proportional to the third moment of the DSD, variance in this data set was consequently higher than in N_t and r_e. Nonetheless, with maximum errors of 0.021 g/m^3 in LWC (profile 1a), 158 cm^{-3} in N_t (profile 1b) and 0.42 µm in r_e (profile 1c) the data allowed for a reasonable interpretation concerning the shapes of the vertical profiles of the microphysical conditions. Apart from overall low values, there was no general conformity between the different profile measurements of this study. Values in Profile 1a rose towards higher altitudes and reached their maximum of 0.14 g/m^3 at approx. 100 m height with small r_e but high N_t values. Above, LWC values never exceeded 0.01 g/m^3. This obviously can be attributed to the fact that the fog's top was exceeded for this part of the profile. Profile 1b reached its maximum of 0.08 g/m^3 already at 32 m and, contrary to 1a, showed less oscillating values within

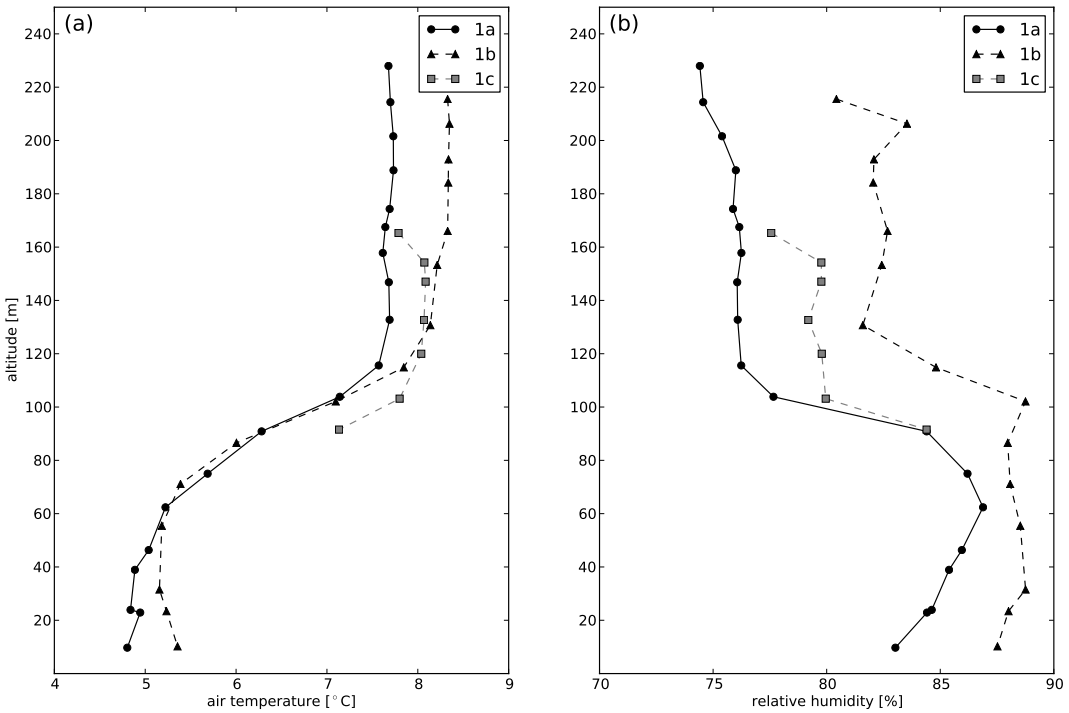

FIGURE 2.4: Profiles of meteorological parameters of October 29, 2011: (a) air temperature, (b) relative humidity

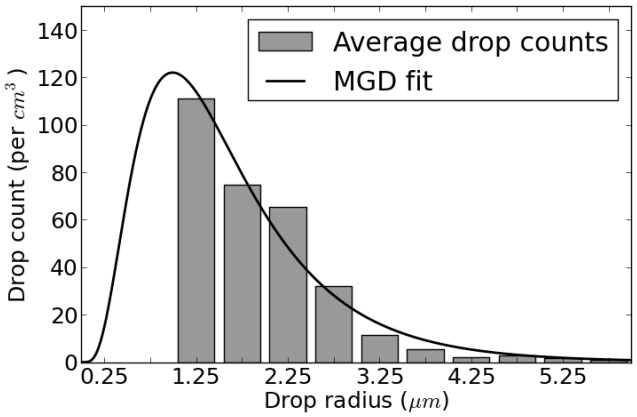

FIGURE 2.5: Average DSD of fog event on October 29, 2011. Bars represent averaged measurements, the curve represents the MGD fit. Only radius sizes up to 6 µm are shown as larger drops did rarely occur.

the fog layer. Towards higher altitudes the values then slowly declined. In profile 1c LWC was reduced close to zero at all altitudes.

The N_t and r_e profiles of Figure 2.6 (b) and 2.6 (c) give further explanations for the shapes of the LWC profiles. MGD-derived N_t values of profiles 1a and 1b showed similar distributions with maxima of 2231 cm^{-3} and 2390 cm^{-3} close to the fog top and second maxima near the ground. r_e values generally rose towards the ground but declined to approx. 2 µm at the very bottom of the profiles. Profile 1b showed a distinct rise in r_e between 25 m and 70 m up to 9.4 µm. While the LWC values of profile 1a were thus mainly driven by total drop counts, in profile 1b the change in drop radii played an equipollent role in the vertical distribution of the DSD, resulting in high LWC values between 20 m and 60 m. In profile 1c drop counts remained close to zero at all measurement heights.

The main condensation layer of profile 1a can be located around 100 m height where high drop counts and small drop radii resulted in large LWC values. Above, rh values indicated a saturation deficit and consequently a rapid decrease of LWC and N_t. Towards the central fog levels, condensation and coalescence of larger particles gained in importance, manifesting themselves in growing r_e and shrinking N_t values. Concurrent value increases of temperature and rh in the lowermost part of profile 1a (up to 60 m) may be ascribed to an increased evaporation of small droplets in these heights caused by the temperature increase. This is reflected in a peak of r_e and a collapse of N_t values at the respective altitudes.

The microphysical parameters of profile 1b indicate a clear vertical development of the fog layer. Above the lower parts (>30 m), clearly increasing LWC and r_e values, as well as decreasing N_t values, point to distinctly stronger droplet growth by coalescence than in the central part of profile 1a. This increase coincided with a change in the temperature profile (see Figure 2.4) from a temperature decrease with height in the lowermost levels towards an isothermal stratification developing into a temperature inversion. Highest fog density is therefore related to the formed inversion base in 1b (uplifted inversion base at 50 m) which typically indicates a shift of the outgoing infra-red radiation maximum at ground level during the measurement of profile 1a (concomitant with a true ground inversion) towards the fog top with maximum cooling and condensation (1b). Above 50 m, temperature increased while r_e and LWC decreased. Although r_e values did not exceed 2.7 µm for the rest of the profile, LWC values only marginally decreased due to significantly higher drop counts towards the inversion base and another peak of N_t above the inversion layer. The slow LWC decline indicates a far-reaching intrusion of the overlying warm air masses into the upper parts of the fog layer which itself induced slow fog dissipation from top to bottom. In profile 1c the dissipation process has already led to N_t and LWC values close to zero at all altitudes.

In summary, the profiles showed considerable changes in their vertical microphysical structure caused by the onset of dissipation. While the LWC values of profile 1a increased with altitude, profile 1b showed an inverted situation with decreasing LWC values towards the fog top, most likely resulting from evaporation and drop settling processes at this stage. In profile 1c the fog's dissipation has already advanced to a complete disappearance of the fog manifesting itself in LWC and N_t values close to zero at all measured altitudes.

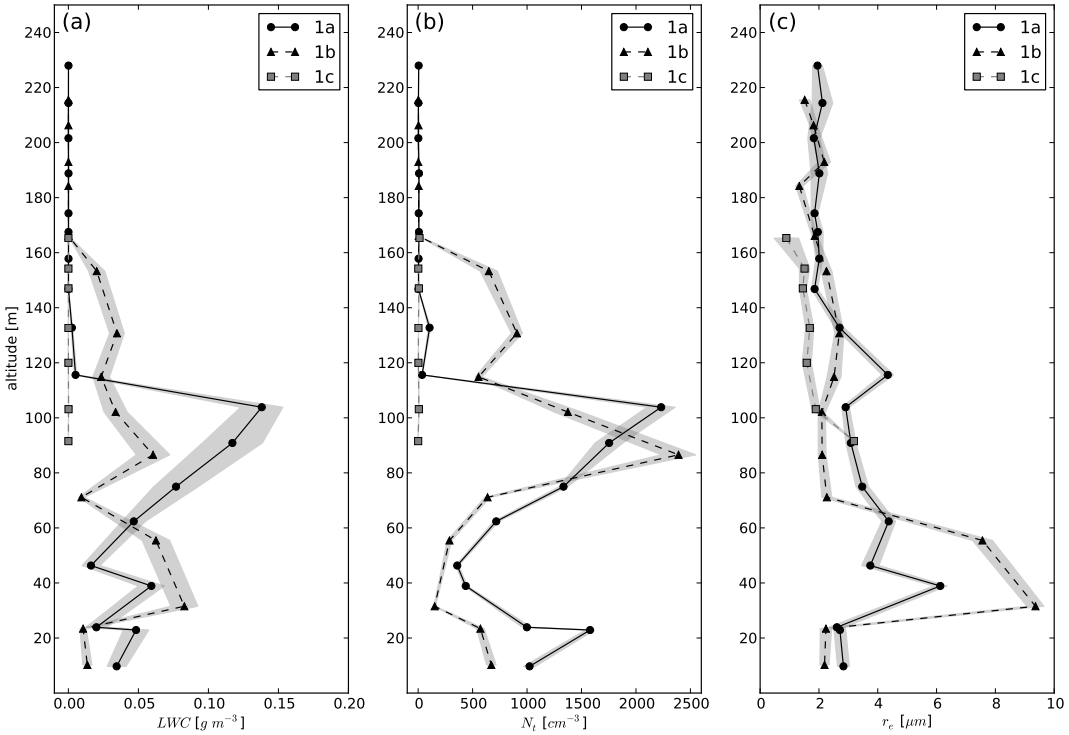

FIGURE 2.6: Profiles of microphysical parameters of October 29, 2011: (a) LWC, (b) N_t, (c) r_e. Grey buffers denote standard errors.

2.3.2 Observations on November 19/20, 2012

On November 19, 2012 the synoptic weather regime over central Europe is described as anticyclonic south-westerly with a well formed high pressure system over Eastern Europe (DWD, 2013b). As a consequence, short-wave troughs approaching from the Atlantic were slowed down over Western Europe and pushed into North-East direction towards Scandinavia. The weather at Linden was thus dominated by the strong high over Eastern Europe throughout the complete measurement period, thus providing adequate preconditions for fog formation (see Figure 2.7).

At 16:21 a dense and long lasting ground fog formed. For a detailed overview of VIS, fog vertical extent and measurement periods see Figure 2.8. Fog/cloud base information was taken from the ceilometer data while fog top altitudes were derived from radar reflectivities. The fog's thickness steadily rose to approx. 200 m until 05:15 on the next morning when it rapidly started increasing in height up to 500 m. The ceilometer records show an oscillating inclination of the cloud base at the very end of the event. After 15 h 53 min this finally resulted in a complete uplift of the fog layer from the ground at 08:13 which formed an LST layer shortly afterwards. The dissipation of the fog layer close to the ground can most likely be ascribed to fog-induced downward terrestrial radiation - observable in a slow temperature rise during the whole event at calm conditions and no change in the synoptic weather situation - as well as additional solar heating after sunrise (06:49). In combination with a continuous rise of the inversion layer and a subsequent lifting of the fog layer itself, these processes lead to dissipation processes at surface level. Before 20:38 no fog top information is depicted as the fog layer did not exceed the near field of the radar device for this period. During the long lasting mature phase (20:27 - 06:23) VIS never exceeded 1 km.

In total, six vertical profiles were recorded - four at the end of the mature phase and

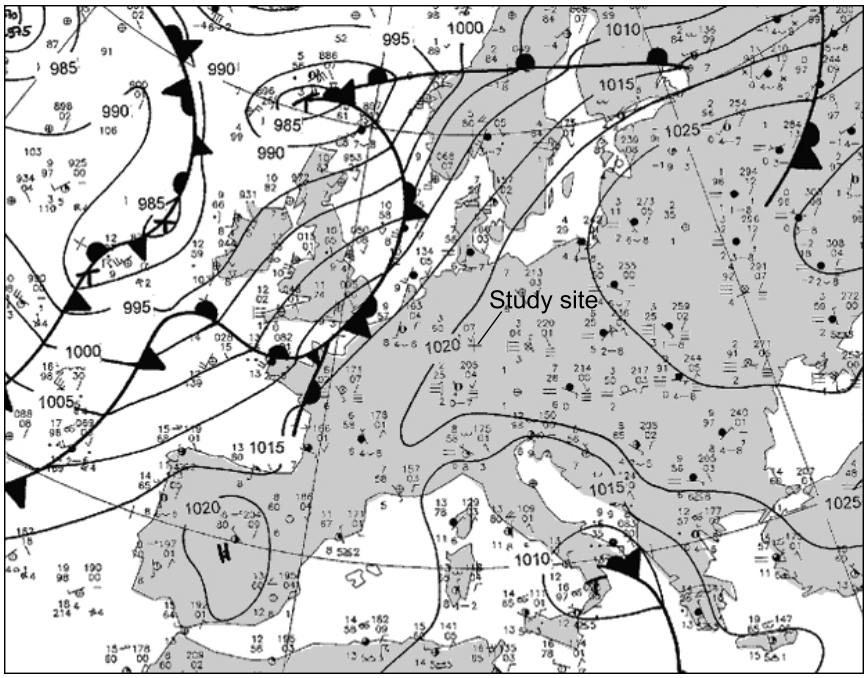

FIGURE 2.7: Synoptic weather situation in Central Europe for November 20, 2012: 00:00 UTC. Modified after DWD (2013a).

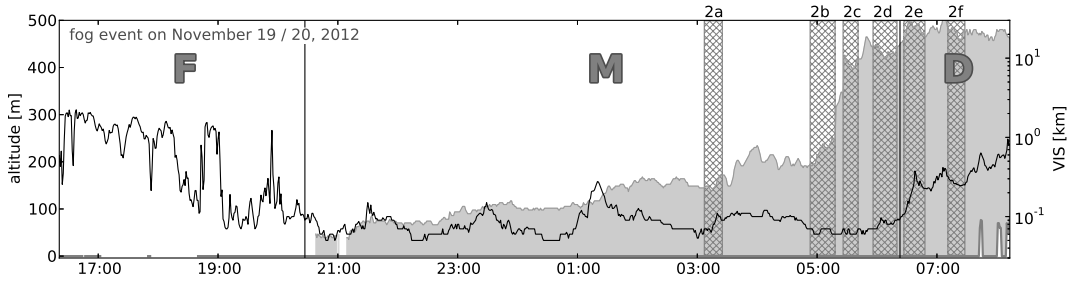

FIGURE 2.8: VIS (black line) & fog vertical extent (grey area) between 16:21 November 19, 2012 and 08:13 November 20, 2012. Fog top altitudes are derived from FMCW radar data (top grey line), cloud base altitudes from ceilometer data (bottom grey line). Hatched stripes mark periods of profile measurements and are labelled corresponding to the following profile plots. Measurement periods were 03:03-03:21 for 2a, 04:49-05:14 for 2b, 05:22-05:37 for 2c, 05:52-06:16 for 2d, 06:23-06:44 for 2e and 07:07-07:24 for 2f. Black vertical lines mark phase transitions. F: Formation, M: Mature fog, D: Dissipation.

two during dissipation. Temperature values of each profile are depicted in Figure 2.9 (a). The profiles are depicted in chronological order, illustrating the general shift towards warmer temperatures during the measurement period. Profile 2a showed a pronounced inversion above 120 m but temperatures slowly started to rise already at 60 m. In profile 2b the inversion layer has been lifted up to altitudes above 140 m with a strong temperature increase starting at 180 m. Below the inversion layer, temperatures decreased with height in 2a and 2b. Profiles 2c to 2f were mainly isothermal and did not show a comparable temperature decrease/increase which can be attributed to a rapid lifting of the inversion layer after 05:15. Mature phase conditions were thus mainly affected by a general temperature decline with minima just below the inversion layer. This is caused by the fog top acting as main emitter of long-wave thermal radiation while the lower fog layers reflect thermal radiation emitted from the surface, consequently feeding the lower parts with additional heat. Fog dissipation, on the other hand, was characterized by isothermal conditions at lower levels. Here, long-wave downward radiation started to cease due to thinning of the fog layer, resulting in less heat adduction close to the surface.

The relative humidity profiles give further insight to the properties of the studied fog event (see Figure 2.9 (b)). Most of the profiles varied between 96 % and 98 % in all altitudes with no substantial outliers. Only profile 2a and 2b showed distinct shifts in the values of the uppermost part. Due to the beginning temperature increase at 60 m in 2a and 140 m in 2b, rh values began to decline, reaching their minima of 91.7 % and 94.5 % at 120 m and 170 m respectively. These heights coincided with rapid temperature increases in both profiles, marking the lower boundary of the inversion layer. Above, rh values rapidly increased towards the profile tops, reaching maxima of 100 % and 97.1 %. These increases can directly be ascribed to the beginning evaporation processes of fog droplets in these altitudes as will be affirmed by the microphysical parameter distributions. Profiles 2c to 2f did not show comparable temperature and rh shifts as they did not reach up to the inversion at the top of the fog layer.

In contrast to the first fog, this event was affected by drop size distributions with much larger but fewer drops (see Figure 2.10). Here, the derived MGD parameters were: $r_c = 2.58$, $\gamma = 1.37$, $a = 11.63$ and $\alpha = 1$. The graph shows moderately distributed drop counts over a wide range of drop size intervals up to radii of 10 µm. This generally resulted in high LWC values at comparably low total drop counts.

Microphysical parameters (LWC, N_t and r_e) of the profiles recorded during mature fog are presented in Figure 2.11. The profiles recorded during dissipation are depicted in Figure 2.12. Compared to the fog event investigated on October 28/29, 2011 the data indicated a significantly denser fog with higher LWC and r_e values but smaller total drop counts. Maximum standard errors were 0.05 g/m^3 in LWC (profile 2b), 54 cm^{-3} in N_t (profile 2f) and 0.72 µm in r_e (profile 2a).

All LWC profiles showed some minor common characteristics but the absolute values differed considerably. In profile 2a LWC values stayed below 0.06 g/m^3 up to 50 m height but changed their average level to 0.1 g/m^3 between 50 m and 120 m altitude. Above 120 m and close to the fog top, the CDP measured rapidly decreasing LWC values. Profile 2b showed considerably larger LWC values throughout the whole measurement but otherwise comparable characteristics. The lower part of the profile had values fluctuating around 0.2 g/m^3, a maximum of 0.44 g/m^3 at 130 m and constantly decreasing values towards the fog top. These value distributions are in good agreement with the meteorological profiles: Maximum LWC values were reached at minimal temperatures and the decline of LWC values with intrusion into the inversion

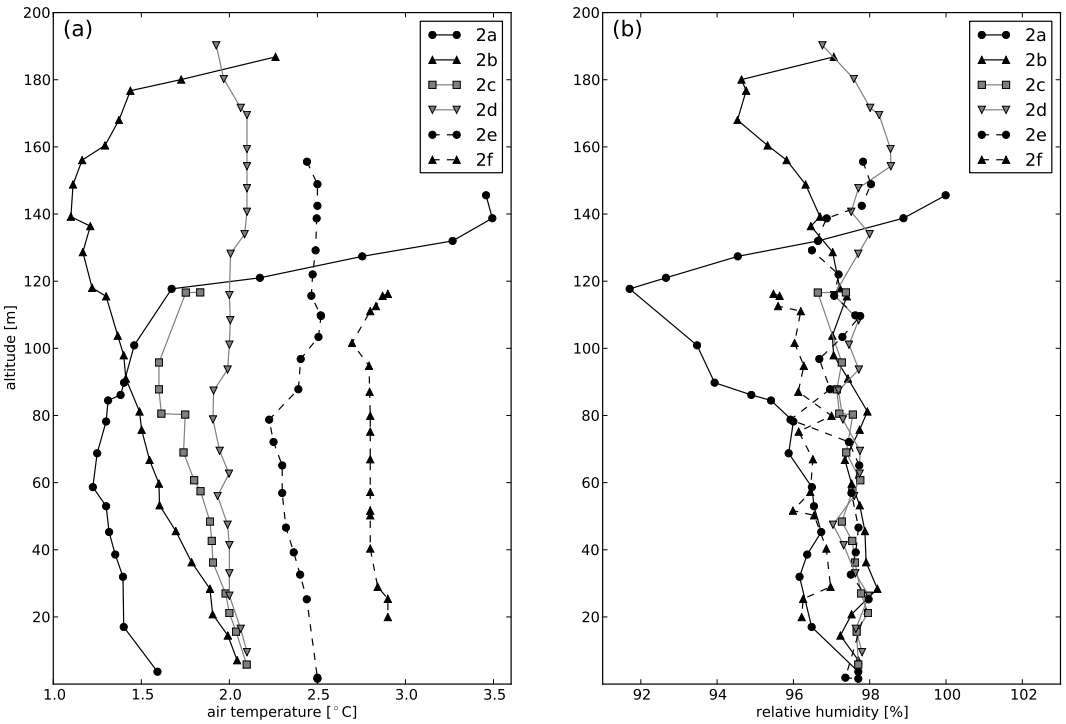

FIGURE 2.9: Profiles of meteorological parameters of November 20, 2012: (a) air temperature, (b) relative humidity

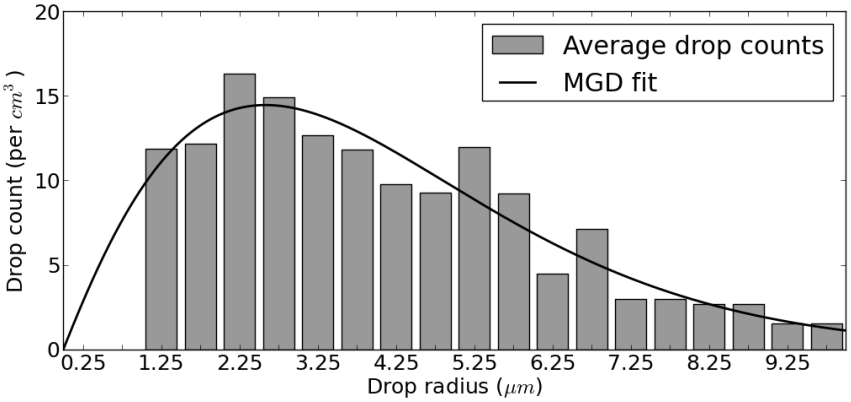

FIGURE 2.10: Average DSD of fog event on November 20, 2012. Bars represent averaged measurements, the curve represents the MGD fit. Only radius sizes up to 10 µm are shown as larger drops did rarely occur.

layer is reflected in growing rh values due to progressive evaporation of fog droplets at these heights.

In profile 2c the curve below 50 m was similarly shaped to 2b with slightly higher LWC values. Above 50 m the values suddenly declined and reached a minimum of 0.04 g/m^3 at 80 m whereas the maximum of this profile was already reached at 116 m with 0.54 g/m^3. It is, however, important to consider that the maximum height during this measurement was also reached at 116 m and thus the profile did not represent the whole fog column as its vertical extent exceeded 400 m at this time (see Figure 2.8). Apart from the strong decrease of LWC between 60 m and 100 m, profile 2c may also have shown similar characteristics to 2a and 2b but due to the missing upper part of the profile it was not possible to derive information for the whole fog column. Surprisingly, in profile 2d LWC values stayed below 0.01 g/m^3 throughout all altitudes although it was recorded during the mature phase. A possible reason for the strong decline in values may be the rapid increase of the fog's vertical extent shortly before, resulting in a smaller density of condensed particles per volume of air. However, a final explanation for this phenomenon could not be determined yet. Profile 2e and 2f showed similar characteristics to 2a although the total fog depth recorded by the radar was considerably higher and thus the profiles did only represent the lower third of the respective fog column.

N_t values (see Figure 2.11 (b)) lay clearly below the values of the October fog event. The shapes of most profiles were similar to the shapes of the corresponding LWC profiles. Maxima of profiles 2a and 2b were measured towards the fog top with 203 and 404 drops per cm^3 and a decline towards higher altitudes. Profile 2c showed highest drop counts with more than 500 drops per cm^3 whereas profile 2d was characterized by values close to zero throughout all altitudes. Profiles 2e and 2f (see Figure 2.12 (b)) showed values comparable to the other profiles for heights above 50 m but considerably smaller values close to the ground. This indicates the emerging uplift of the fog layer from the surface during the measurement of these profiles as a consequence of the onset of dissipation at that time.

All recorded r_e profiles were characterized by very little vertical variability but considerable shifts between the different profiles. For profiles 2c to 2f no vertical trends could be identified. Only profile 2a and 2b showed a slight increase in the central region and a distinctive decrease towards the respective profile tops. In conjunction with decreasing N_t values, shrinking drop sizes were thus jointly responsible for the diminishing LWC values at the fog top. Due to rising temperatures within the inversion layer, the fog top layer during mature stage was characterized by strong evaporation processes affecting both, small and large drops (decrease of N_t due to dissipation of small drops and decrease of r_e due to beginning evaporation of large drops).

In summary, all presented parameters showed changes in the microphysical characteristics within small temporal and vertical intervals. The shapes of the recorded N_t profiles closely matched the curves of the respective LWC profiles whereas r_e values did not show noticeable similarities except for profile 2a and 2b. Vertical changes in LWC within one profile could thus mainly be ascribed to changes in the distribution of total drop counts and not to changes in their size. On the other hand, the apparent differences of r_e values between the profiles led to the conclusion that drop growth and shrinking were responsible for long term temporal changes in LWC over the whole fog column. However, phase-dependent profile characteristics were hard to detect in this fog event as they were superimposed by the short term changes that manifested themselves primarily in shifts of r_e values throughout the whole fog column.

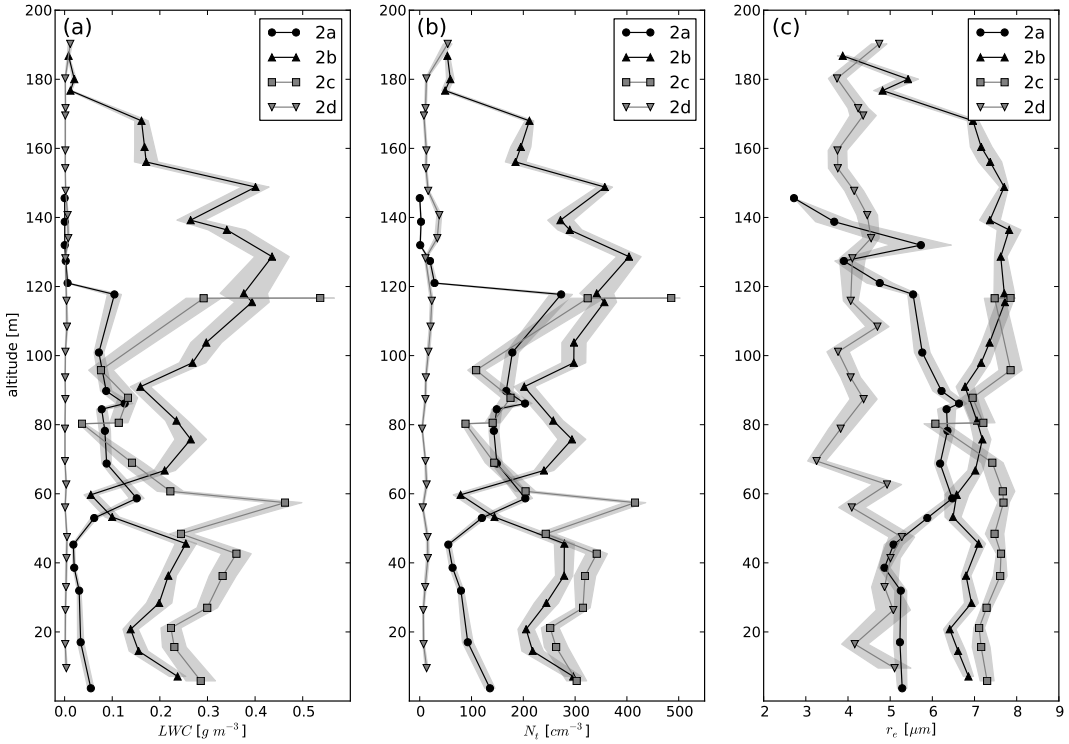

FIGURE 2.11: Profiles of microphysical parameters of November 20, 2012 (Mature fog): (a) LWC, (b) N_t, (c) r_e. Grey buffers denote standard errors.

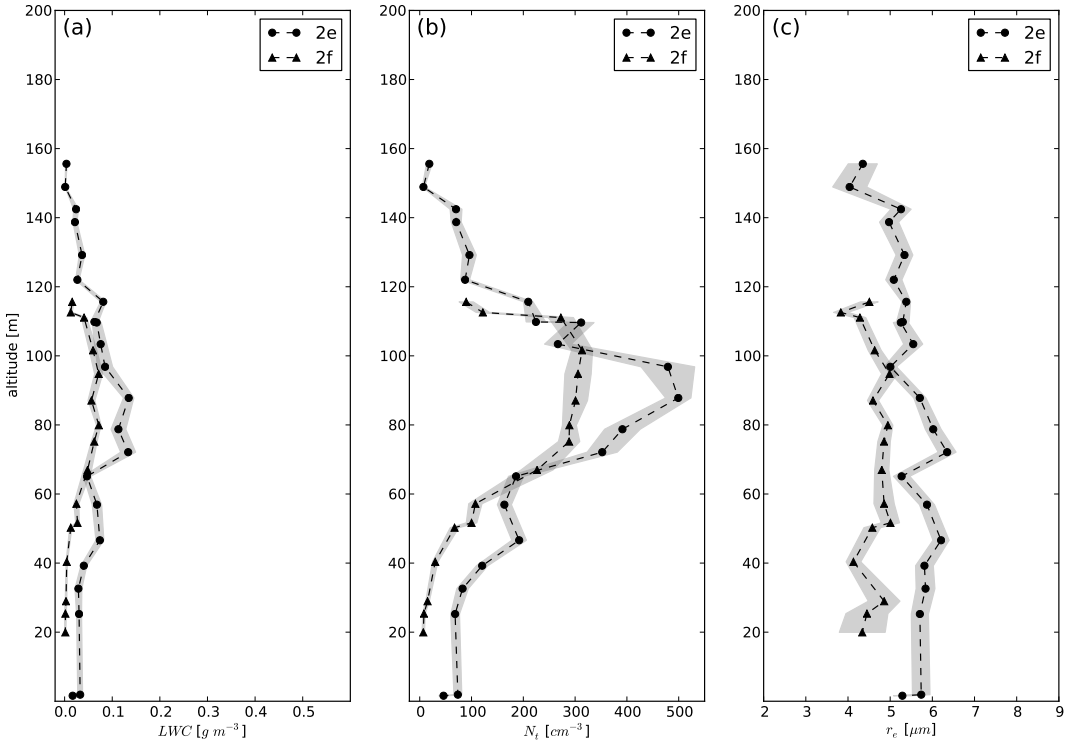

FIGURE 2.12: Profiles of microphysical parameters of November 20, 2012 (Dissipation): (a) LWC, (b) N_t, (c) r_e. Grey buffers denote standard errors.

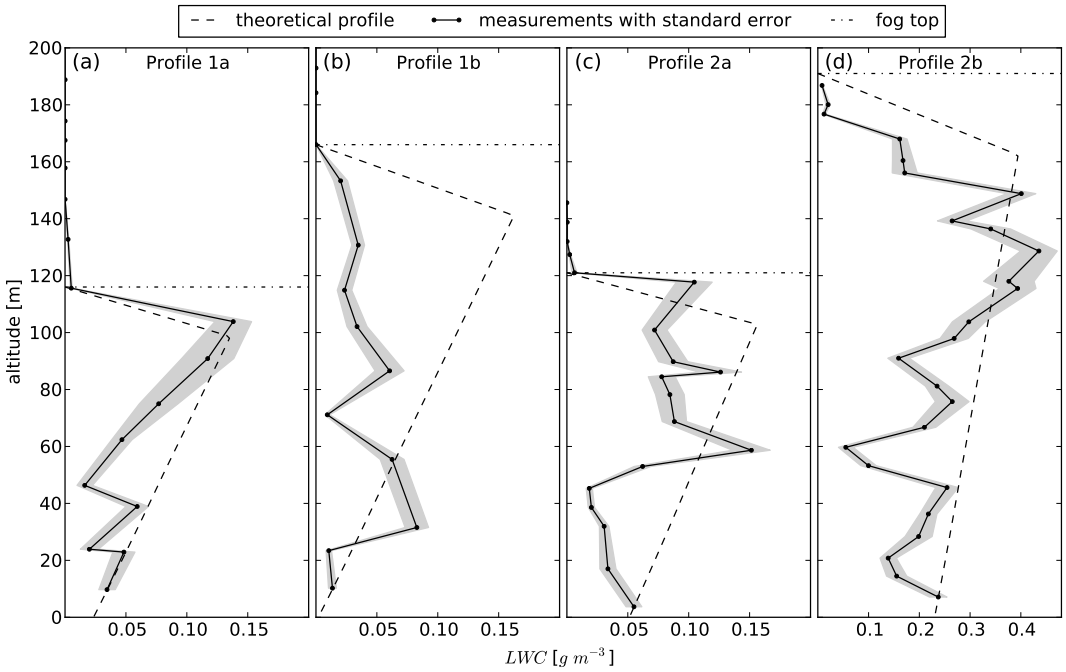

FIGURE 2.13: Comparison between measured (solid) and theoretical (dashed) LWC profiles. Grey buffers denote standard errors. (a) and (b) show profiles 1a and 1b of October 29, 2011; (c) and (d) show profiles 2a and 2b of November 20, 2012. Profiles 1a, 2a and 2b were recorded during the mature phase whereas profile 1b was recorded during dissipation. Fog top altitudes are indicated by the thin dashed horizontal line. All profiles are depicted together with their theoretical equivalent.

2.3.3 Comparison Between Theoretical Profiles and Measurements

To investigate the validity of the theoretical assumptions about the vertical LWC distribution made by Cermak and Bendix (2011), suitable profiles, covering the whole fog column, were compared to their theoretical equivalent. Only four out of all measured profiles did meet these requirements (coverage of the whole fog vertical extent with continuous records throughout the whole fog column) and are discussed in the following section. The rest of the profiles either did not reach into the fog layer (1c) because the measurement had to be aborted or they were recorded during a period when the fog layer did significantly outreach the maximum ascent level of 250 m (2c to 2f). Fog top altitude information was derived for each profile by deriving VIS values from measured drop size distributions via Equation 2.7.

Figure 2.13 depicts the investigated profiles together with their theoretical equivalents. Profile 1a (mature phase) showed broad agreement between LWC measurements and theoretical expectations, leading to a mean absolute error (MAE) of 0.024 g/m^3 and a difference of 21.2 % in the derived liquid water paths (LWP, defined as the integrated LWC over the whole fog column in g/m^2). Apart from some minor deviations in the central part, shape and maximum of the theoretical profile were closely resembled by the CDP records. However, the distribution of N_t and r_e values (see Figure 2.6 (b) and (c)) showed, that increases in LWC values could mainly be ascribed to increases of drop counts and not to particle growth.

In contrary, the measurements of profile 1b (dissipation phase) did not match the theoretical distribution (Figure 2.13 (b)). The LWC profile rather showed an inverted

shape with slowly decreasing values towards the fog top and a fast increase at the bottom. This led to LWC values that persisted at lower values compared to the theoretical profile, especially at higher altitudes. An MAE of 0.055 g/m^3 and an LWP difference of 57.3 % supported this perception. While N_t values followed a similar distribution compared to those of profile 1a, r_e values were generally smaller except for the range between 25 m and 70 m where they were higher, causing the reported increase of LWC values at the bottom of the fog. A possible explanation for the difference between theoretical and measured profile can be found in the onset of the dissipation phase shortly before the measurement of this profile. The Smart Tether data support the assumption, that mixing and subsequent evaporation processes from above have already set in at this point. These processes may then have led to the slowly sinking LWC values towards the fog top and high values at the bottom due to settling of the remaining larger drops.

In profile 2a (Figure 2.13 (c)) (mature phase) LWC measurements followed the theoretical expectations with increasing values towards higher altitudes. Apart from the similar trend, there was a distinct vertical variance in the data leading to differences between measured and theoretical profile. This resulted in an MAE of 0.042 g/m^3 and an LWP difference of 32.8 %. While r_e values in profile 1a and 1b played a minor role in profile-internal LWC changes, here drop size changes apparently contributed to the shape of the LWC profile. In short, r_e rose slowly with height and fell rapidly at the fog top. Together with similarly shaped N_t distributions this resulted in the described LWC profile.

Profile 2b (Figure 2.13 (d)) (mature phase) showed the highest LWC values during this case study. With an MAE of 0.086 g/m^3 it also showed the highest absolute difference to its theoretical counterpart. However, the relative LWP difference was only at 25.6 % which is the second best value of the four investigated profiles. Again, the measured LWC values showed high variance but generally followed the theoretical distribution. Similarly to profile 2a, drop size changes influenced the LWC distribution to some extent. However, the LWC profile was basically governed by the shape of the corresponding N_t profile.

Referring to the theoretical LWC model of Cermak and Bendix (2011), a good agreement in trend and gradient between measured and theoretical LWC profiles was found for the profiles measured during the mature stage. For the LWC profile measured during the dissipation stage less accordance with the theoretical profile had to be stated. Summarized LWP properties, given in Table 2.2, also point to the conjecture that minimal accordance between theoretical and measured profiles can be found for profiles with minimal LWP values (see profile 1b with 5.829 g/m^2 and a Δ_{LWP} of 57.3 %). In most profiles, LWC did slowly increase with height and none of the profiles showed a lapse rate higher than the sub-adiabatic postulation. Rapid decreases of LWC values close to the fog top were also found in most of the records. However, profile-internal changes of LWC were mainly governed by changes in N_t whereas r_e was relatively constant at all altitudes. This indicates, that the increase of drop counts due to condensation at remaining nuclei had a greater influence on the fog's density formation than drop growth by coalescence processes. The expected increase in drop size towards greater altitudes as well as constant drop counts in the central fog layer as presented in Brenguier et al. (2000) and Chang and Heston (2002) for low stratiform clouds could thus not be affirmed for the observed radiation fog events. Concerning the anticipated shift from adiabatic to sub-adiabatic conditions close to the ground (Cermak and Bendix, 2011), it was not possible to come to a meaningful conclusion.

A comparison with the measurements from the literature showed that both, Okita (1962) and Pinnick et al. (1978), found generally larger LWC values, than one would

Profile	LWP_{meas}	LWP_{theo}	MAE	abs. Δ_{LWP}	rel. Δ_{LWP}
1a	7.135	9.053	0.024	1.918	21.2 %
1b	5.829	13.659	0.055	7.830	57.3 %
2a	8.162	12.140	0.042	3.977	32.8 %
2b	42.056	56.520	0.086	14.464	25.6 %
Ø	15.796	22.843	0.052	7.047	34.2 %
Okita	40.950	36.941	0.091	-3.757	−12.7 %
Pinnick	56.038	34.633	0.097	-16.322	−27.9 %

TABLE 2.2: Measured and theoretical LWP, mean absolute error (MAE) as well as absolute and relative LWP differences compared to the theoretical profile's LWP of 1a, 1b, 2a, 2b and averages from Okita (1962) and Pinnick et al. (1978).

have expected based on the postulated sub-adiabatic lapse rate (see Table 2.2). On average, the records of Okita (1962) resulted in an MAE of 0.091 g/m^3 and an LWP difference of −12.7 %. However, LWC distributions showed the same general trend and N_t majorly determined the shape of the LWC profile whereas mean volume radii showed only a slight negative slope towards higher altitudes. The measurements of Pinnick et al. (1978) showed less accordance with the theoretical profile leading to an MAE of 0.097 g/m^3 and an LWP difference of −27.9 %. For a given LWP this would result in an overestimation of the fog's geometrical thickness. However, a direct comparison between the different data sets is difficult due to the different instrumentations used. All studies, however, are congruent in affirming the general rise of LWC values and none of the studies were able to reinforce the assumption of constant drop counts and growing drop sizes with altitude. On the contrary, r_e was found to be relatively constant within one profile and N_t showed strong variation as well as growing numbers towards the fog top.

2.4 Conclusion

The aim of the present study was to verify the validity of a theoretical fog LWC profile and its applicability for satellite based ground fog detection with a special consideration of the temporal evolution during the fog life cycle. For this purpose LWC profiles were measured during two different fog events by means of a tethered balloon borne composition of meteorological and microphysical measurement instruments including an optical particle counter.

The measurements indicate a good agreement in trend and gradient between measured and theoretical LWC profiles for mature stage observations. The LWC profile measured during the dissipation stage shows less accordance with the theoretical LWC profile. Although the proposed vertical function of Cermak and Bendix (2011) was in general agreement with the LWC measurements during the mature phase, there were some discrepancies due to the vertical variance in the measured profiles. A distinct temporal variability in drop size distributions provided for additional uncertainties in the profile comparison. Especially during the onset of dissipation, microphysical conditions were subject to changes on small spatial and temporal scales and because of the small size of the data set an extended analysis with a larger number of profiles was not possible. As fog dissipation can also be subject to a variety of processes, one should not expect a single representative profile to cover all possibilities. This is why local thermodynamics and meteorological conditions should be accounted for during the derivation of a dissipation profile. In this way, different processes, e.g. dissipation

without compensational flows from above or turbulent synoptic-advective processes could also be represented in the model.

However, the available data indicated that the theoretical considerations concerning the vertical distribution of microphysical fog features should be adapted. The different phases of the fog life cycle should be considered when formulating and adapting the theoretical LWC profile used in satellite based ground fog detection methods. With respect to the poor accordance between measurement and theory in profiles with small LWP values, these methods could also be adapted by accounting for different theoretical LWC profiles that are assigned to different LWP ranges. LWP derivations from satellite data (e.g. Kawamoto and Nakajima (2001)) could then be used as an adjustment in the vertical LWC function for the calculation of the fog's geometrical thickness. However, for a well-founded adaptation of the theoretical LWC profile under consideration of the fog's evolutionary stages and the observed LWP distributions, more LWC profiles during a great variety of fog events have to be measured.

Accordingly, there is a growing interest for an operational measurement method in addition to the balloon-borne system concerning a comprehensive data set of LWC profiles. Such a measurement method could be realized using the 94 GHz FMCW radar introduced in this study. This radar system is highly suitable for fog detection and monitoring (Bennett et al., 2009; Maier et al., 2013). The measurement system presented in this study can be used to develop a proper relationship between the radar reflectivity (Z) and the liquid water content that could be used to invert measured Z into LWC values.

Acknowledgements

The authors thank the German Research Foundation DFG for funding the project (BE1780/14-1; TH1531/1-1). The work is also part of the COST action EG Climet. The authors are also thankful to Sebastian Achilles for his support during the measurements.

Bibliography

Anasphere (2012). *SmartTether Operating Manual*. Anasphere Inc., Bozeman, MT, USA, 2 edition.

Bendix, J. (2002). Working Group I , National Report Germany. Technical report, COST 722.

Bendix, J., Eugster, W., and Klemm, O. (2011). Fog – boon or bane? *Erdkunde*, 65(3):229–232.

Bendix, J., Thies, B., Cermak, J., and Nauß, T. (2005). Ground Fog Detection from Space Based on MODIS Daytime Data—A Feasibility Study. *Weather and Forecasting*, 20(6):989–1005.

Bendix, J., Thies, B., Nauß, T., and Cermak, J. (2006). A feasibility study of daytime fog and low stratus detection with TERRA/AQUA-MODIS over land. *Meteorological Applications*, 13(2):111–125.

Bennett, A., Nash, J., Gaffard, C., Moyna, B., Oldfield, M., and Huggard, P. (2009). Observations from the UK Met Office 94 GHz FMCW cloud radar. In *Proceedings of*

the 8th International Symposium on Tropospheric Profiling, number October, pages 2–5, Delft.

Bergot, T., Terradellas, E., Cuxart, J., Mira, A., Leichti, O., Mueller, M., and Nielsen, N. W. (2007). Intercomparison of single-column numerical models for the prediction of radiation fog. *Journal of Applied Meteorology and Climatology*, 46(4):504–521.

Betts, A. K. (1982). Cloud thermodynamic models in saturation point coordinates. *Journal of the Atmospheric Sciences*, 39(10):2182–2191.

Biral (2010). *Operation and Maintenance Manual for the VPF-730 Present Weather Sensor & VPF-710 Visibility Sensor*. Biral, Portishead, Bristol, UK.

Biral (2012). *HSS VPF -730 Combined Visibility & Present Weather Sensor*. Bristol Industrial and Research Associates Limited, Portishead, Bristol, UK.

Boers, R. and Mitchell, R. (1994). Absorption feedback in stratocumulus clouds influence on cloud top albedo. *Tellus A*, 46(0280-6495):229–241.

Bott, A. and Trautmann, T. (2002). PAFOG - a new efficient forecast model of radiation fog and low-level stratiform clouds. *Atmospheric Research*, 64(1-4):191–203.

Brenguier, J.-L., Pawlowska, H., Schüller, L., Preusker, R., Fischer, J., and Fouquart, Y. (2000). Radiative Properties of Boundary Layer Clouds: Droplet Effective Radius versus Number Concentration. *Journal of the Atmospheric Sciences*, 57(6):803–821.

Bruijnzeel, S., Eugster, W., and Burkard, R. (2005). Fog as a Hydrologic Input. In *Encyclopedia of Hydrological Sciences*, pages 559–582. John Wiley & Sons, Ltd, Chichester, UK.

Cermak, J. (2006). *SOFOS - A new Satellite-based Operational Fog Observation Scheme*. PhD thesis, Philipps-University of Marburg.

Cermak, J. and Bendix, J. (2007). Dynamical nighttime fog/low stratus detection based on Meteosat SEVIRI data: A feasibility study. *Pure and Applied Geophysics*, 164(6-7):1179–1192.

Cermak, J. and Bendix, J. (2011). Detecting ground fog from space – a microphysics-based approach. *International Journal of Remote Sensing*, 32(12):3345–3371.

Chang, S. S. and Heston, W. D. W. (2002). The clinical role of prostate-specific membrane antigen (PSMA). *Urologic Oncology*, 7(1):7–12.

Choularton, T. W., Fullarton, G., Latham, J., Mill, C. S., Smith, M. H., and Stromberg, I. M. (1981). A field study of radiation fog in meppen, West Germany. *Quarterly Journal of the Royal Meteorological Society*, 107(452):381–394.

Danne, O. (1996). *Messungen physikalischer Eigenschaften stratiformer Bewölkung mit einem 94 GHz - Wolkenradar*. Phd thesis, University of Hannover.

Deirmendjian, D. (1969). *Electromagnetic Scattering on Spherical Polydispersions*. American Elsevier Publishing Company, New York, USA.

DMT (2009). *Data Analysis User's Guide Chapter I : Single Particle Light Scattering*. Droplet Measurement Technologies, Boulder, CO, USA.

DMT (2010). *Cloud Droplet Probe (CDP) Manual*. Droplet Measurement Technologies, Boulder, CO, USA.

DWD (2013a). Bodenanalysekarten Europa.

DWD (2013b). Die Grosswetterlagen Europas.

Ellrod, G. P. (1995). Advances in the Detection and Analysis of Fog at Night Using GOES Multispectral Infrared Imagery. *Weather and Forecasting*, 10(3):606–619.

Eyre, J. R., Brownscombe, J. L., and Allam, R. J. (1984). Detection of fog at night using Advanced Very High Resolution Radiometer (AVHRR) imagery. *Meteorological Magazine*, 113(1346):266–271.

Fuzzi, S., Facchini, M. C., Orsi, G., Lind, J. A., Wobrock, W., Kessel, M., Maser, R., Jaeschke, W., Enderle, K. H., Arends, B. G., Berner, A., Solly, I., Kruisz, C., Reischl, G., Pahl, S., Kaminski, U., Winkler, P., Ogren, J. A., Noone, K. J., Hallberg, A., Fierlinger-oberlinninger, H., Puxbaum, H., Marzorati, A., Hansson, H.-C., Wiedensohler, A., Svenningsson, I. B., Martinsson, B. G., Schell, D., and Georgii, H. W. (1992). The Po Valley Fog Experiment 1989. *Tellus B*, 44(5):448–468.

Fuzzi, S., Laj, P., Ricci, L., Orsi, G., Heintzenberg, J., Wendisch, M., Yuskiewicz, B., Mertes, S., Orsini, D., Schwanz, M., Wiedensohler, A., Stratmann, F., Berg, O. H., Swietlicki, E., Frank, G., Martinsson, B. G., Günther, A., Dierssen, J. P., Schell, D., Jaeschke, W., Berner, A., Dusek, U., Galambos, Z., Kruisz, C., Mesfin, N. S., Wobrock, W., Arends, B., and Ten Brink, H. (1998). Overview of the Po valley fog experiment 1994 (CHEMDROP). *Contributions to atmospheric physics*, 71(1):3–19.

Garbow, B. S., Hillstrom, K. E., and More, J. J. (1980). Documentation for MINPACK subroutine LMDIF.

Glickman, T. S. (2000). *Glossary of Meteorology (2nd ed.)*. American Meteorological Society, Boston, 2 edition.

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of integrals, sums, series, and products*. Elsevier, 7 edition.

Gultepe, I., Müller, M. D., and Boybeyi, Z. (2006). A new visibility parameterization for warm-fog applications in numerical weather prediction models. *Journal of Applied Meteorology and Climatology*, 45(11):1469–1480.

Gultepe, I., Tardif, R., Michaelides, S. C., Cermak, J., Bott, A., Bendix, J., Müller, M. D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., and Cober, S. G. (2007). Fog Research: A Review of Past Achievements and Future Perspectives. *Pure and Applied Geophysics*, 164(6-7):1121–1159.

Hayasaka, T., Nakajima, T., Fujiyoshi, Y., Ishizaka, Y., Takeda, T., and Tanaka, M. (1995). Geometrical thickness, liquid water content, and radiative properties of stratocumulus clouds over the western North Pacific. *Journal of Applied Meteorology*, 34(2):460–470.

Hess, P. and Brezowsky, H. (1977). *Katalog der Grosswetterlagen Europas 1881–1976, 3. verbesserte und ergänzte Aufl.* Deutscher Wetterdienst, Offenbach am Main, Germany.

Huggard, P. G., Oldfield, M. L., Moyna, B. P., Ellison, B. N., Matheson, D. N., Bennett, A. J., Gaffard, C., Oakley, T., and Nash, J. (2008). 94 GHz FMCW cloud radar. *Proceedings of SPIE*, 7117:711704–711704–6.

Hutchison, K. D. (2002). The retrieval of cloud base heights from MODIS and three-dimensional cloud fields from NASA's EOS Aqua mission. *International Journal of Remote Sensing*, 23(24):5249–5265.

Iwabuchi, H. and Hayasaka, T. (2002). Effects of Cloud Horizontal Inhomogeneity on the Optical Thickness Retrieved from Moderate-Resolution Satellite Data. *Journal of the Atmospheric Sciences*, 59(14):2227–2242.

Jacobs, W., Nietosvaara, V., Bott, A., Bendix, J., Cermak, J., Michaelides, S., and Gultepe, I. (2008). *Short range forecasting methods of fog , visibility and low clouds - COST 722*. Office for Official Publications of the European Communities.

James, P. M. (2007). An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe. *Theoretical and Applied Climatology*, 88(1-2):17–42.

Kawamoto, K. and Nakajima, T. (2001). A global determination of cloud microphysics with AVHRR remote sensing. *Journal of Climate*, 14(9):2054–2068.

Liu, D., Yang, J., Niu, S., and Li, Z. (2011). On the evolution and structure of a radiation fog event in Nanjing. *Advances in Atmospheric Sciences*, 28(1):223–237.

Maier, F., Bendix, J., and Thies, B. (2012). Simulating Z-LWC relations in natural fogs with radiative transfer calculations for future application to a cloud radar profiler. *Pure and Applied Geophysics*, 169(5-6):793–807.

Maier, F., Bendix, J., and Thies, B. (2013). Development and application of a method for the objective differentiation of fog life cycle phases. *Tellus, Series B: Chemical and Physical Meteorology*, 65(1):1–17.

Minnis, P., Young, D. F., Kratz, D. P., Coakley, J. A., King, M. D., Garber, D. P., Heck, P. W., Mayor, S., and Arduini, R. F. (1997). Cloud optical property retrieval (subsystem 4.3). *Clouds and the Earth's Radiant Energy System (CERES) Algorithm Theoretical Basis Document*, pages 1–60.

Nemery, B., Hoet, P. H. M., and Nemmar, A. (2001). The Meuse Valley fog of 1930: An air pollution disaster. *Lancet*, 357(9257):704–708.

Okita, T. (1962). Observations of the vertical structure of a stratus cloud and radiation fogs in relation to the mechanism of drizzle formation. *Tellus*, 14(3):310–322.

Pagowski, M., Gultepe, I., and King, P. (2004). Analysis and Modeling of an Extremely Dense Fog Event in Southern Ontario. *Journal of Applied Meteorology*, 43(1):3–16.

PHC (2010). *T2-03 Micro Diaphragm Pumps Data Sheet*. PHC Precision Fluidics Division, Hollis, NH, USA.

Pilié, R., Eddie, W., Mack, E., Rogers, C., and Kocmond, W. (1972). PROJECT FOG DROPS Part I : Investigations of Warm Fog Properties. Technical Report August, Cornell Aeronautical Laboratory, Inc., Buffalo, N.Y.

Pilié, R. J., Mack, E. J., Kocmond, W. C., Eadie, W. J., and Rogers, C. W. (1975a). The Life Cycle of Valley Fog. Part II: Fog Microphysics. *Journal of Applied Meteorology*, 14(3):364–374.

Pilié, R. J., Mack, E. J., Kocmond, W. C., Rogers, C. W., and Eadie, W. J. (1975b). The Life Cycle of Valley Fog. Part I: Micrometeorological Characteristics. *Journal of Applied Meteorology*, 14(3):347–363.

Pinnick, R. G., Hoihjelle, D. L., Fernandez, G., Stenmark, E. B., Lindberg, J. D., Hoidale, G. B., and Jennings, S. G. (1978). Vertical Structure in Atmospheric Fog and Haze and Its Effects on Visible and Infrared Extinction. *Journal of the Atmospheric Sciences*, 35(10):2020–2032.

Pinto, R., Larrain, H., Cereceda, P., Lázaro, P., Osses, P., and S., S. R. (2001). Monitoring fog-vegetation communities at a fog-site in Alto Patache, South of Iquique, Northern Chile, during "El Niño" and "La Niña" events (1997-2000). In *Second International Conference on Fog and Fog Collection*, pages 293–296, Ottawa, Canada. International Development Research Center.

Pruppacher, H. R., Klett, J. D., and Wang, P. K. (1998). *Microphysics of Clouds and Precipitation*, volume 28. Kluwer Academic Publisher, Doordrecht, Netherlands, 2 edition.

Roach, W. T., Brown, R., Caughey, S. J., Crease, B. a., and Slingo, A. (1982). A field study of nocturnal stratocumulus: I. Mean structure and budgets. *Quarterly Journal of the Royal Meteorological Society*, 108(455):103–123.

Ronda, R. J., Steeneveld, G. J., and Holtslag, A. A. M. (2011). Can we forecast radiation fog using mesoscale models? Technical Report October 2005, Wageningen University, Wageningen.

Schulze-Neuhoff, H. (1976). Nebelfeinanalyse mittels zusätzlicher 420 Klimastationen - Taktische Analyse 1:2 statt 1:5 Mill. *Meteorologische Rundschau*, 29(3):75–84.

Slingo, A., Nicholls, S., and Schmetz, J. (1982). Aircraft observations of marine stratocumulus during JASIN. *Quarterly Journal of the Royal Meteorological Society*, 108(458):833–856.

Spiegel, J. K., Zieger, P., Bukowiecki, N., Hammer, E., Weingartner, E., and Eugster, W. (2012). Evaluating the capabilities and uncertainties of droplet measurements for the fog droplet spectrometer (FM-100). *Atmospheric Measurement Techniques*, 5(9):2237–2260.

Tampieri, F. and Tomasi, C. (1976). Size distribution models of stratospheric particles in terms of the modified gamma function. *Archiv für Meteorologie, Geophysik und Bioklimatologie Serie A*, 25(1):47–54.

Tardif, R. (2007). The impact of vertical resolution in the explicit numerical forecasting of radiation fog: A case study. *Pure and Applied Geophysics*, 164(6-7):1221–1240.

Terradellas, E. and Bergot, T. (2007). Comparison between two-single-column models designed for short-terms fog and low-clouds forecasting. *Física de la Tierra*, pages 189–203.

Thoma, C., Schneider, W., Masbou, M., and Bott, A. (2012). Integration of local observations into the one dimensional fog model PAFOG. *Pure and Applied Geophysics*, 169(5-6):881–893.

Turner, J., Allam, R. J., and Maine, D. R. (1986). A case study of the detection of fog at night using channels 3 and 4 on the Advanced Very High Resolution Radiometer (AVHRR). *Meteorological Magazine*, 115:285–290.

Vaisala (2012). *CL31 Ceilometer for Cloud Height Detection Data sheet*. Vaisala, Helsinki, Finland.

van de Hulst, H. C. (2012). *Light scattering by small particles*. Courier Dover Publications, Mineola N.Y.

Wang, J., Daum, P. H., Yum, S. S., Liu, Y., Senum, G. I., Lu, M. L., Seinfeld, J. H., and Jonsson, H. (2009). Observations of marine stratocumulus microphysics and implications for processes controlling droplet spectra: Results from the Marine Stratus/Stratocumulus Experiment. *Journal of Geophysical Research Atmospheres*, 114(18):D18210.

Wendisch, M., Mertes, S., Heintzenberg, J., Wiedensohler, A., Schell, D., Wobrock, W., Frank, G., Martinsson, B. G., Fuzzi, S., Orsi, G., Kos, G., and Berner, A. (1998). Drop size distribution and LWC in Po Valley fog. *Contributions to atmospheric physics*, 71:87–100.

Wieprecht, W., Acker, K., Mertes, S., Collett, J., Jaeschke, W., Brüggemann, E., Möller, D., and Herrmann, H. (2005). Cloud physics and cloud water sampler comparison during FEBUKO. *Atmospheric Environment*, 39(23-24):4267–4277.

Wobrock, W., Schell, D., Maser, R., Kessel, M., Jaeschke, W., Fuzzi, S., Facchini, M. C., Orsi, G., Marzorati, a., Winkler, P., Arends, B. G., and Bendix, J. (1992). Meteorological characteristics of the Po Valley fog. *Tellus B*, 44(5):469–488.

Chapter 3

A 10 Year Fog and Low Stratus Climatology for Europe Based on Meteosat Second Generation Data

This chapter has been published as: Egli, S., Thies, B., Drönner, J., Cermak, J., & Bendix, J. (2017). A 10 year fog and low stratus climatology for Europe based on Meteosat Second Generation data. Quarterly Journal of the Royal Meteorological Society, 143(702), 530–541.

Received: 10 May 2016 / Accepted: 06 October 2016
https://doi.org/10.1002/qj.2941

Abstract

Due to the significant impact of fog and low stratus (FLS) on economy, ecology and traffic systems, there is a growing demand for high resolution information on FLS occurrence. In this study, a baseline climatology of FLS $\mathrm{h\,d^{-1}}$ based on data recorded from 2006 to 2015 by the Spinning Enhanced Visible and Infrared Imager system aboard the Meteosat Second Generation satellites (SEVIRI) is computed for Europe to provide the requested information. It is the first 10 year, spatially explicit climatology for FLS based on data with a temporal resolution of 15 minutes. The data set is validated against Meteorological Aviation Routine Weather Reports (METAR) and shows good accordance with an average Heidke Skill Score of 0.45. Temporal and spatial variations in FLS frequency as well as interannual trends are analyzed. Winter season shows highest FLS occurrence but a general decrease over the investigated period. Spring, summer and fall show less pronounced trends and lower average FLS frequencies. Possible reasons for these distributions are discussed.

3.1 Introduction

Fog and low stratus (FLS) are of considerable importance to a variety of features in climatic processes and ecological systems. FLS often acts as a water supply to otherwise arid ecosystems (Bruijnzeel et al., 2005; Von Glasow and Bott, 1999) and by influencing radiative transfer processes in the atmospheric boundary layer it acts as a modifier in the climate system (Vautard et al., 2009).

FLS also directly interferes with many aspects of human life. By reducing visibility it often leads to hazardous situations in air, sea and road traffic (Bendix et al., 2011).

Due to modifications in the heat fluxes of the lower atmosphere, FLS also prolongs inverted weather situations, thus often reducing air quality significantly (Nemery et al., 2001).

Fog is internationally defined as a suspension of small droplets in the air, reducing visibility to less than 1000 m at the surface (Glickman, 2000). As for many applications, e.g. air traffic and radiative transfer processes, the difference between fog and low stratus is of minor importance (Cermak et al., 2009), during this study both are investigated collectively.

Due to the vast number of influence factors of FLS there is a growing demand for area-wide, spatially high resolution data on FLS occurrence and trends in its distribution. A long term, spatially explicit FLS climatology is needed to provide the requested information. Currently, no suitable information is available. Several studies have been using synoptic observations to derive spatial fog patterns and to investigate long term trends in FLS distribution (e.g. Giulianelli et al., 2014; Sachweh and Koepke, 1995, 1997; Schulze-Neuhoff, 1976; Vautard et al., 2009). However, ground-based measurements of visibility and cloud cover are often sparsely spread. Additionally, different temporal resolutions, variations in quality control and discontinuities in the time series recorded by meteorological stations in different countries complicate the interpretation of the data and make it difficult to derive spatiotemporally continuous information reliably for larger areas. Other schemes like the one presented in Scherrer and Appenzeller (2014) use indirect approaches via sunshine duration measurements to derive climatological FLS series but they can only give valid results for single locations.

A meaningful FLS climatology that is representative on a continental scale should be based on a homogeneously derived data set with high resolutions both in space and time. Records from polar orbiting satellite systems like NOAA or MODIS provide high spatial resolutions but suffer from poor temporal resolution. Although FLS frequencies have been investigated using polar orbiting systems (cf. Bendix and Bachmann, 1991; Bendix, 2002; Musial et al., 2014) geostationary satellite platforms (GEO) are more suitable to provide continuous information over long time periods with a high temporal resolution while assuring a frequent measurement interval.

An attempt to detect FLS on the basis of GEO satellite data was conducted by Güls and Bendix (1996) using the WEFAX Product derived from Meteosat First Generation (MFG) data. However, due to the poor spectral resolution of the MFG sensor, the resulting FLS masks were only of limited use, especially for patchy FLS structures. The only long term investigation of FLS occurrence in Europe on the basis of GEO satellite data with an adequate spatiotemporal and spectral resolution was done by Cermak et al. (2009) for the years between 2004 and 2008 using Meteosat Second Generation (MSG) data. FLS was detected by applying the day- and nighttime approaches developed in Cermak and Bendix (2007) and Cermak and Bendix (2008). However, data was only processed for the winter months (December to February) when FLS occurrences reach their highest values in Europe.

In this study, the schemes developed by Cermak and Bendix (2007) and Cermak and Bendix (2008) have been adapted to allow for shorter processing times while additionally some of the subtests used in these approaches have been modified to account for varying conditions in different domain regions and for seasonal changes. Data acquired by the geostationary MSG platform between 2006 and 2015 have been used to derive FLS patterns on a representative data basis for all seasons on the whole 24 hour cycle. The focus of this study has been put on Europe. The results are validated against METAR data recorded during the same period. FLS frequency

distributions and trends as well as the dependence on its meteorological drivers are analyzed in detail.

At this point, it should be stressed that the FLS product generated in this study does not give any information about FLS base heights. A distinction between low laying stratus layers and ground fog is thus not possible based on this data set. However, the generated FLS data are the basis for the calculation of a true ground fog product as shown in Cermak and Bendix (2011).

In section 3.2 satellite and ground truth data acquired for the investigation as well as the applied processing and validation schemes are presented. Validation results of the FLS product are given in section 3.3. The resulting baseline FLS climatology is discussed in section 3.4 with final conclusions given in section 3.5.

3.2 Data and Methods

3.2.1 Geostationary Satellite Data

As a basis to this study, Meteosat Second Generation data from the platforms of Meteosat-8, -9 and -10 have been used. Meteosat-8 was placed over 3.4°W/0°N while Meteosat-9 and Meteosat-10 were situated over 0°E/0°N during their operational usage. The on-board Spinning Enhanced Visible and Infrared Imager system (SEVIRI) scans the full hemisphere every 15 min with a sub-satellite resolution of 3 km in 11 spectral bands and an extra HRV channel with a resolution of 1 km as opposed to only 3 bands on the previously operated MFG systems (Aminou et al., 1997). The electromagnetic spectrum covered reaches from visible wavelengths (0.56 µm) to thermal infrared (14.40 µm) which allows for a more detailed examination of the atmospheric column in these wavelengths (Schmetz et al., 2002). All recorded scenes between 2006 and 2015 were acquired from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) in HRIT Level 1.5 format. These data are geolocated and corrected for all radiometric and geometric effects, making them suitable for the derivation of meteorological products (EUMETSAT, 2013).

3.2.2 FLS Detection Scheme

FLS is considered to be a low stratiform cloud layer with small droplets in the liquid phase not exceeding sizes of 20 µm (Best, 1951; Cermak, 2006; Egli et al., 2015; Roach, 1995). The FLS detection scheme applied in this survey uses a number of threshold tests to filter pixels from MSG records that match these criteria (Figure 3.1).

First, all cloud pixels are identified using a combination of the 10.8 µm and 3.9 µm channels. This approach allows for a distinction between cloud free and cloud contaminated pixels due to the uniformly low reflectivity of clear sky pixels in both bands while cloudy pixels generally show higher intensities at 3.9 µm than at 10.8 µm. In a second step, snow pixels are excluded from the original mask using the Normalized Difference Snow Index (NDSI) introduced by Dozier (1989). A subsequent cloud phase test uses 10.8 µm brightness temperatures (T_B) (cf. Robert A. Houze, 1993) in combination with differences in the absorption characteristics of liquid water and ice in the 12.0 µm and the 8.7 µm channels to remove all non liquid clouds (Cermak, 2006). A small droplet proxy test filters out all large sized cloud drops for daytime scenes by making use of different emissive characteristics of small and large sized cloud droplets in the middle infrared around 3.9 µm. Stratiformity of the clouds top surface is assumed when standard deviations of the brightness temperature in the 10.8 µm channel are not exceeding 2.0 K. A plausibility test checks the remaining cloud layers

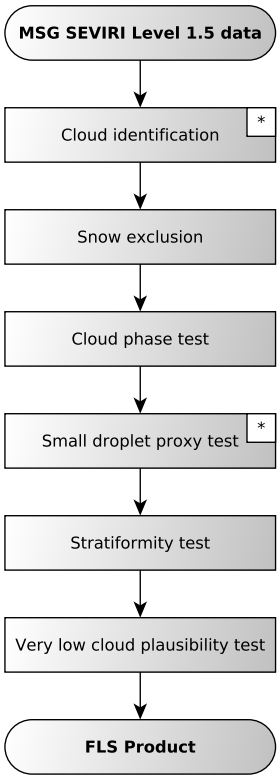

FIGURE 3.1: FLS detection scheme (simplified) as developed by Cermak (2006). Tests marked with a * were adapted for this study.

for their height by comparing brightness temperature differences of each cloud with the surrounding clear sky pixels, thus reducing the final mask to low level stratus and fog layers only. All tests are explained in detail in Cermak (2006) and Cermak and Bendix (2007, 2008).

During the operational application of the FLS scheme, some improvements were done to the original tests. Particularly, cloud identification as well as small droplet filtering have been adapted to allow for a dynamic domain selection and a more precise threshold determination. To speed up processing times, all tests have been parallelized additionally using the OpenCL standard maintained by the Khronos Group (Stone et al., 2010).

Adaptation of the Cloud Identification

The cloud identification approach as developed by Cermak (2006) uses a global threshold in the 10.8 µm and 3.9 µm band difference to distinguish FLS pixels from other cloud types and clear sky pixels. This threshold is dynamically computed for each MSG record and is then applied to the complete scene. Depending on the size of the domain, this can lead to inaccurate threshold values because changes in the geographic location of the pixels lead to changes in the satellite's viewing angle and the sun zenith angle, which in turn influence the radiation intensity perceived at the satellite sensor. Furthermore, different surface types cause shifts in spectral band differences even if cloud conditions do not change and thus should not be treated equally.

To avoid these inaccuracies, the algorithm was adapted to feature dynamical pixel-wise thresholds (Figure 3.2). To consider different surface types and different illumination conditions during the computation process, threshold values are calculated

separately for land and sea areas as well as for day- and nighttime. The land/sea mask is derived from the WorldClim DEM (Hijmans et al., 2005). The day/night mask is based on the sun zenith angle of the respective pixel. To account for shifts in the satellite's viewing angle, a moving window technique is applied. The satellite scene is traversed by a window of 450 by 450 pixels size in steps of 50 pixels (Figure 3.2 (a)). To avoid edge effects along domain borders, a buffer of 200 pixels is included in the moving window procedure. The subsequent histogram analysis as described in Cermak and Bendix (2007) for nighttime and Cermak and Bendix (2008) for daytime scenes is processed separately for every step and each mask combination.

Small droplets show smaller emissivity differences between 10.8 µm and 3.9 µm, than both, large droplet clouds and cloud free areas (Hunt, 1973). This leads to a strong right-shift of pixels with predominantly small droplets in the difference histograms at night. Other cloud types with larger drop sizes do not show this effect. Differences between the 10.8 µm and the 3.9 µm band are higher for these clouds than for small droplet clouds and cloud-free areas. This results in a strong left-shift of pixels covered by large droplet clouds (see Figure 3.2 (a) night histograms). Consequently, for nighttime, upper and lower thresholds have to be computed in order to distinguish cloud-covered pixels from cloud-free areas.

During daytime, solar contamination causes high signals in the 3.9 µm band for all cloud types due to their high reflective properties. Cloud-free areas show generally lower reflectances. In the difference histograms, this leads to a distinct left-shift of all cloud covered pixels in comparison to cloud-free areas (see Figure 3.2 (a) day histograms). Thus, for daytime only lower thresholds have to be calculated in order to capture all cloud covered pixels.

To avoid sharp edges at each 50 pixel step interval in the cloud identification output, derived threshold values of each step are assigned to window centers only (Figure 3.2 (b)). As the reliability of the histogram approach is sensible to changes in the value distribution, outliers exceeding two standard deviations of their 3 by 3 neighborhood are eliminated by replacing them with the mean of their neighbors. The area between window centers is then filled using bilinear interpolation (Figure 3.2 (c)). This is done separately for both, lower and upper threshold values. Window and step sizes are empirically adjusted to ascertain an adequate representation of local conditions while still guaranteeing a sufficient pixel count suitable for the histogram approach as well as an adequate consideration of the continuous characteristics of atmosphere and cloud cover. Following this procedure, derived thresholds allow for a distinction between cloud free and cloud contaminated pixels at day and an additional separation into small and large droplet clouds at night. The final cloud mask is computed by applying the interpolated threshold values to the band difference of the 10.8 µm and the 3.9 µm channel.

Adaptation of the Small Droplet Proxy Test

The small droplet proxy test introduced by Cermak (2006) is used to separate small and large droplets at daytime. This is needed as the cloud identification approach can only provide a binary classification mask (cloud / no cloud) for daytime scenes due to solar signal contamination in the 3.9 µm channel. It can, however, not distinguish between large and small drops, the latter of which are predominant in fogs and low stratus clouds, often not exceeding radii of 20 µm (Best, 1951; Choularton et al., 1981; Egli et al., 2015; Pilié et al., 1975; Roach, 1995).

Thus, in this test, large droplet clouds are distinguished from small droplet clouds by making use of the radiative properties of FLS in the middle infrared (3.9 µm). Small

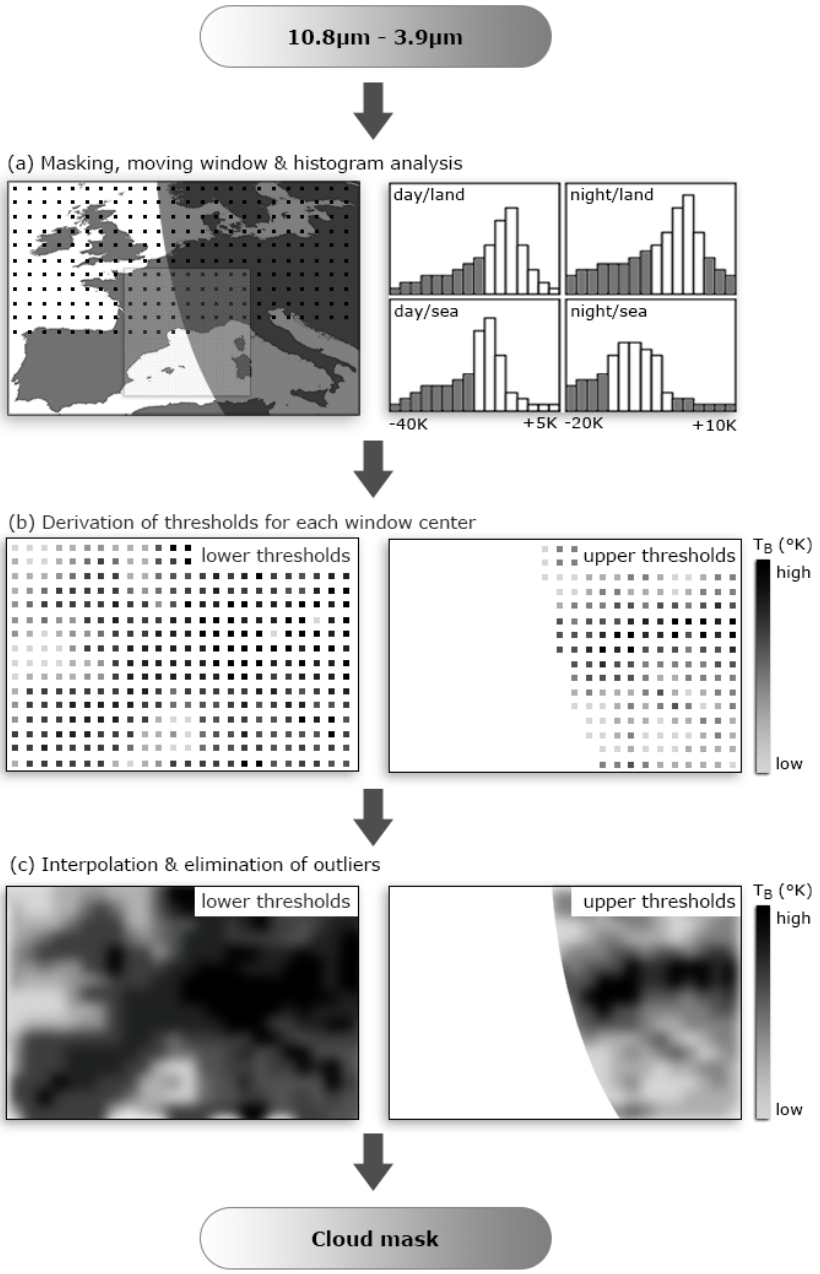

FIGURE 3.2: Processing scheme of the cloud identification algorithm. (a) Left: Calculation of land/sea & day/night masks. The 450 by 450 pixel window is highlighted for one step exemplarily. Black dots mark already processed steps. Right: Histogram analysis done for each step with one histogram for every mask combination. Grey bars denote values interpreted as "cloudy". (b) Derived threshold values for each step of the moving window procedure. Lower thresholds are computed for all mask combinations, upper thresholds only for night mask combinations. (c) Elimination of outliers and interpolation of threshold values for final cloud mask.

drops, predominant in FLS, are reflective at these wavelengths and cause a high solar component in the signal received at the satellite platform. This causes the total signal to exceed that of large droplet clouds as well as cloud free areas. On the other hand, large droplet clouds generally show low thermal signals around 3.9 µm due to their altitude and because of their weak reflectivity properties in these wavelengths. This causes the satellite to perceive a weaker signal for large droplet clouds than for cloud free areas (Cermak and Bendix, 2008). Conclusively, as FLS shows higher values than cloud free regions and clouds containing large droplets have lower values than cloud free regions in the 3.9 µm band, both (FLS and large droplet clouds) can be separated using the signal perceived from cloud free areas. By averaging clear pixel values in the 3.9 µm band adjacent to cloudy areas and using these values as thresholds, the small droplet proxy test distinguishes areas with small droplet clouds from areas with large droplet clouds during daylight. The solar contamination of the 3.9 µm band thus plays an important role in the distinction between large and small droplet areas and it is therefore not being removed from the signal.

The 3.9 µm channel partly overlaps with the CO_2 absorption band centered around 4.2 µm. As CO_2 absorption varies depending on the satellite zenith angle, in this study the small droplet proxy test was adapted to allow for a dynamic threshold computation (Cermak and Bendix, 2007). This was achieved by using a similar approach to the cloud threshold calculation described in chapter 3.2.2 and by following the same moving window technique. For each step, the average of the radiance values of the 3.9 µm band is calculated for all cloud free land-pixels within the respective window area. Just as in the cloud identification scheme, these averages are then assigned to window centers and get interpolated to all remaining pixels to produce more fine grained thresholds. In the next step, pixels that were classified as cloudy by the cloud identification test are sub-classified into small and large droplet clouds using these detected thresholds.

3.2.3 Validation

For validation purposes a vast ground truth data set was acquired covering the complete period from 2006 to 2015. When comparing satellite derived and ground truth data, several fundamental insufficiencies have to be considered. According to Cermak and Bendix (2008) and Obregon et al. (2014) subpixel effects make it difficult to assign a pixel to a category (FLS yes/no) for a specific point in time, especially at the margins of FLS patches where only parts of the pixels are covered. Slight instabilities in the satellite's orbit and perspective as well as parallax shifts cause collocation issues which lead to ambiguities in the assignment of ground truth data locations to pixel coordinates. As the SEVIRI system needs about 12 min for a full disk scan (Aminou, 2002) and because of different temporal resolutions in both datasets, timing uncertainties arise which complicate data comparison additionally. Another problem is caused by multiple cloud layers, especially when one of the layers completely obscures the sky. In this case, satellite and meteorological station records coercively result in different observations.

Meteorological stations reporting surface synoptic observations (SYNOP) are well spread over Europe and data is easily accessible. However, report intervals range between 3 h and 6 h only which makes SYNOP data not suitable for a direct validation of the FLS product with its 15 min temporal resolution.

Meteorological Aviation Routine Weather Reports (METAR) recorded by airport stations are similarly well distributed in Europe. With a report interval between 20 min and 60 min their information can be compared to the FLS product almost

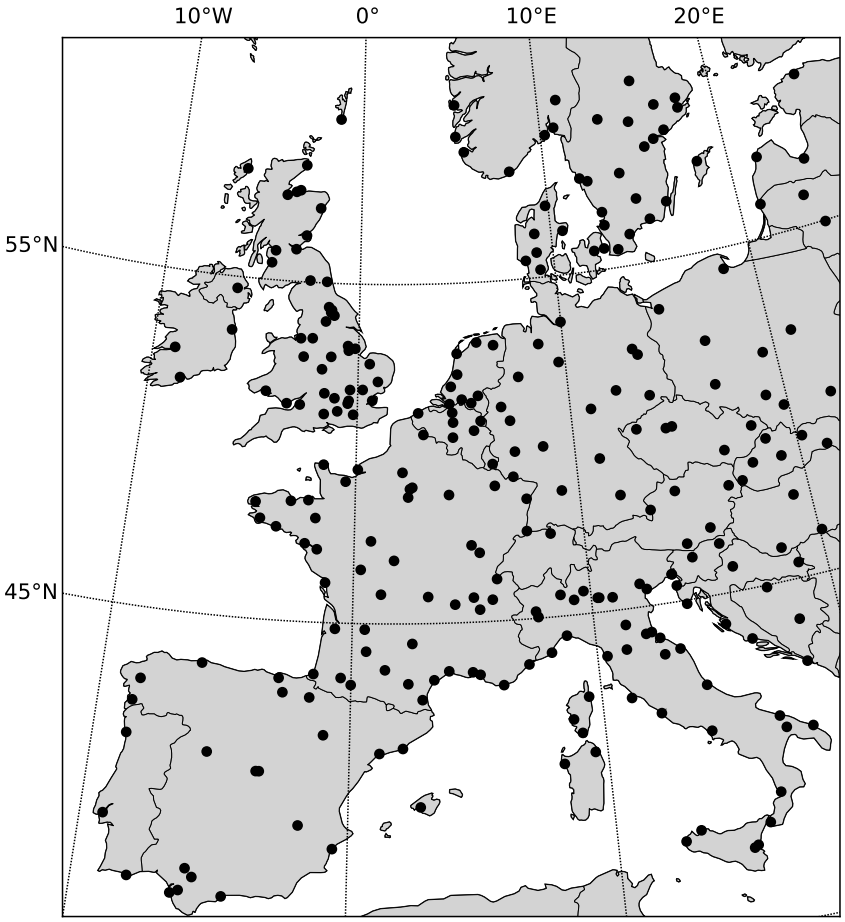

FIGURE 3.3: All 273 METAR stations used for validation (black dots). Only land stations with intervals smaller than 30 min and a 24 h report cycle covering the complete period from 2006 to 2015 were taken into account.

directly. However, data quality varies between single METAR stations as measurement instruments and techniques are not standardized. While most big airports detect horizontal visibility (VIS), cloud cover (CC) and cloud height (CH) using ceilometer and visibility sensors, some smaller airfields rely on simple human observation only. Additionally, report times and intervals as well as temporal coverage vary between METAR stations. To ensure comparability between the FLS product and the METAR data set, only land stations with intervals smaller than 30 min and a 24 h report cycle covering the complete period from 2006 to 2015 were taken into account. In this way, the temporal offset between METAR reports and FLS product did never exceed 15 min. After filtering, 273 METAR reporting stations were included in the final validation process (Figure 3.3).

METARs represent observations from the earth's perspective whereas the FLS product is derived from information recorded from a satellite's point of view. As described above, a straightforward comparison between both datasets is problematic. However, METAR information can still be used as a proxy for FLS occurrence as perceived by the MSG system. Whenever VIS is reported to decrease below 1000 m or the lowest cloud layer reaches less than 1000 m altitude, FLS is assumed. In the second case, CC must also be exceeding 5/8 of the sky (CC in METAR is given in oktas). For validation, all interpretable METAR data were merged with the corresponding FLS

data using a simple temporal nearest neighbour technique. To account for collocation issues and subpixel effects neighbouring pixels were included in the data aggregation as described in Cermak (2006) and Cermak and Bendix (2008). On the basis of the resulting data set the following verification measures were computed.

$$\mathrm{PC} = \frac{\mathrm{a} + \mathrm{d}}{\mathrm{a} + \mathrm{b} + \mathrm{c} + \mathrm{d}} \tag{3.1}$$

$$\mathrm{POD} = \frac{\mathrm{a}}{\mathrm{a} + \mathrm{c}} \tag{3.2}$$

$$\mathrm{FAR} = \frac{\mathrm{b}}{\mathrm{a} + \mathrm{b}} \tag{3.3}$$

where PC is Percentage Correct, POD is Probability Of Detection and FAR is False Alarm Ratio. a is the sum of all correctly predicted FLS occurrences (True Positive) while b is the sum of all incorrectly predicted FLS occurrences (False Positive). c is the sum of False Negatives and d is the sum of True Negatives accordingly.

These measures address very specific characteristics of the data quality only. Thus, additionally Heidke Skill Scores (HSS) were calculated which compare the forecast accuracy of a product to the forecast accuracy of a random guess procedure, thereby giving an overall indication of the validity of the algorithm output.

$$\mathrm{HSS} = \frac{2(\mathrm{ad} - \mathrm{bc})}{(\mathrm{a} + \mathrm{c})(\mathrm{c} + \mathrm{d}) + (\mathrm{a} + \mathrm{b})(\mathrm{b} + \mathrm{d})} \tag{3.4}$$

As described in section 3.2.2 and 3.2.2, inaccuracies in the FLS detection routine due to changes in the satellite's viewing angle were accounted for by using a dynamic moving window approach. During the validation process, the quality measures were thus also calculated for each METAR reporting station separately. In this way, it was possible to identify a potential bias in the FLS data as a result of the geographic location as well as changing satellite viewing angles.

3.3 Validation Results

All validation measures were computed for all MSG scenes that were processed, covering the complete period between 2006 and 2015. The values of a, b, c and d are given for each month separately in Table 3.1 in the supporting information material. Relative counts of monthly FLS observations made by METAR stations are plotted in Figure 3.4 together with corresponding monthly averages of each validation measure.

FLS observations as well as validation measures show a distinct seasonal trend. According to METAR records, most FLS situations occur during winter season between November and February with a maximum of $6.3\,\mathrm{h\,d^{-1}}$ in January. On the other hand, FLS on average only lasts for $2.0\,\mathrm{h\,d^{-1}}$ between the six months from April till September with only $1.7\,\mathrm{h\,d^{-1}}$ in July. Interannual variability is very low in all months with a maximum standard deviation of only $0.98\,\mathrm{h\,d^{-1}}$ in November.

PC shows best results during summer month with values peaking in July (93.23 %) whereas in winter season PC values decline to a minimum of 78.32 % in January. However, high values of PC can be ascribed to low counts of FLS observations during summer season which in turn lead to a high count of true negatives (correctly detected non-FLS situations) but not a higher count of true positives coercively. With standard deviations not exceeding 3.24 % the variability between the years is very low.

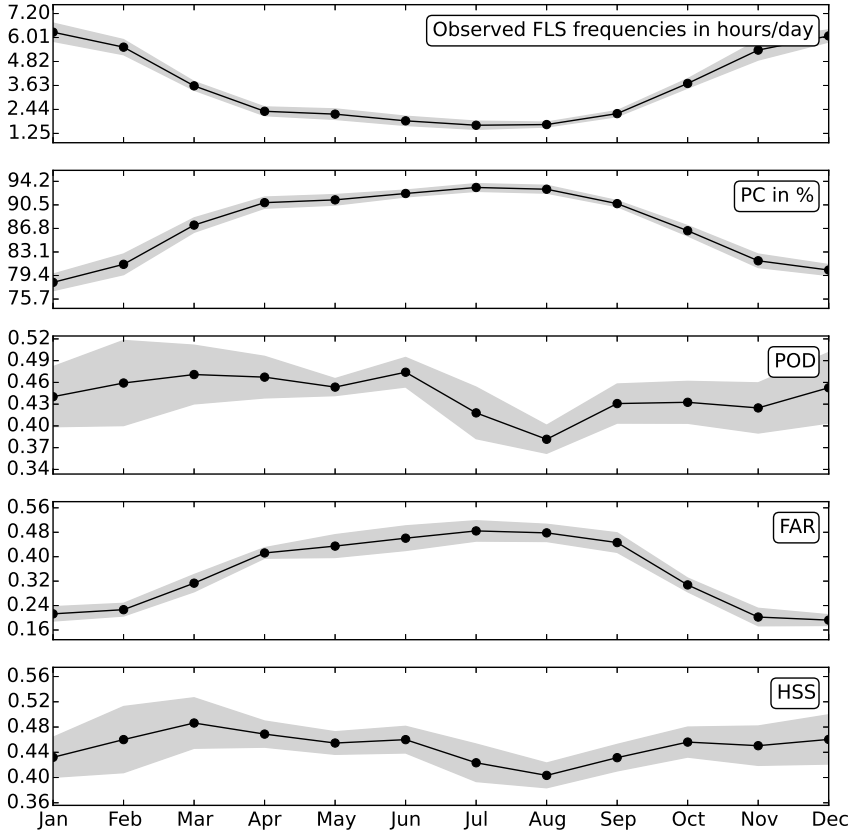

FIGURE 3.4: Validation results. Monthly averages of PC, POD, FAR, and HSS as well as observed FLS frequencies in $\mathrm{h\,d^{-1}}$ averaged over all reporting METAR stations. Gray areas denote standard deviations.

POD is constantly on a high level around 0.46 during the first half of the year. After a maximum in June with 0.47 POD rapidly drops to a minimum of 0.38 in August but recovers till the end of the year. Therefore the probability for a correct hit is slightly higher in winter and spring than during summer and fall. The interannual variability is relatively high with 0.12 in February but generally smaller during the summer months with a minimum of 0.02 in May. Information about prediction accuracy of Non-FLS situations, however, is not represented by the POD.

FAR values show the same trend as PC and similarly low variability between the years with a maximum standard deviation of 0.08 in June. Maxima of FAR are reached during summer with 0.48 in July whereas winter months are characterized by lower values for FAR with 0.19 in December. This distribution emphasizes that during winter less FLS incidences are falsely predicted whereas in summer misses are more often. However, as FLS decreases in the summer months, the remaining few FLS occurrences are more likely to be missed and thus a higher FAR is an expected outcome.

HSS values and standard deviations are similarly distributed as the POD curves with a peak in March (0.49), a minimum in August (0.40) and generally lower standard deviations in summer than in winter. Winter and spring show good performances whereas July, August and September suffer from slightly worse output quality as a result of coinciding high FAR and low POD values during these months. However, the overall algorithm performance is not severely influenced by the seasonal cycle with values never falling below 0.40.

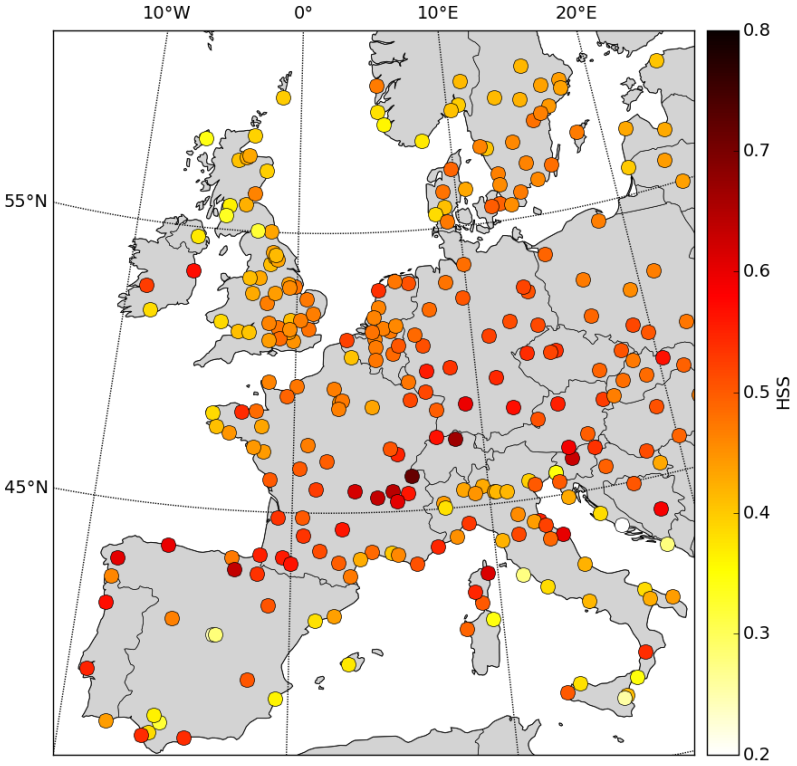

FIGURE 3.5: Dependence of data validity on geographic location. Depicted are HSS values of all 273 METAR stations averaged over the complete 10 year period from 2006 to 2015.

The data's quality dependence on the geographic location is depicted in Figure 3.5. Stations close to coastlines, especially in the Mediterranean area, are subject to lower HSS values whereas stations in continental regions show overall better validation results. Besides variations in the general quality of METAR records between different stations, one possible reason for this distribution are changes in the predominating FLS type depending on coastal distance. FLS emergence at the coast is most often due to advection of FLS layers from the sea which is related to unstable and windy conditions. Inland FLS frequency is more likely a consequence of a high count of radiative FLS events during stable anticyclonic and calm wind conditions. As these conditions favor single cloud layer situations and due to the slow movement of the FLS layers, the detection scheme is more likely to detect the location of an FLS layer precisely which consequently makes it easier for the scheme to distinguish FLS from non FLS pixels. Another reason for the decline of HSS values at coastal regions can be ascribed to the sudden change in ground surface texture and albedo as well as complex topographic conditions along coastlines. This leads to mixed pixel effects that make it difficult to derive precise thresholds for these areas which is especially true for the complex structure of the northern Adriatic coastline where the FLS detection scheme shows its weakest performance. A special treatment of coast pixels similar to the approach used in the improved APOLLO scheme (Saunders and Kriebel, 1988) that is applied to the Advanced Very High Resolution Radiometer (AVHRR) could yield better results for coastal areas. However, as both, cloud identification and small droplet proxy test derive thresholds based on areal information of a 450 by 450 pixel window, a simple pixel based extra treatment of coast pixels is not possible when using the histogram approach that was applied in this study.

Summarizing, the validation results show good quality of the investigated FLS data set. FAR and POD suffer from slight drawbacks during summer, partly as a result of decreasing FLS occurrences during this season. Although HSS values are dependent on the geographic location with better results in continental areas than along coastlines, generally HSS performance is convincing throughout the year. Again, it has to be stressed, that subpixel effects, collocation issues, temporal shifts between both datasets and multiple cloud layers cause difficulties when comparing satellite derived area wide information with local ground based measurements. As no preliminary filtering was applied to the data to avoid obvious misclassifications (no removal of scenes with high clouds overlaying FLS layers, no averaging of multiple scenes, no removal of twilight scenes, consideration of all seasons) and because of the high temporal resolution of the data set of only 15 min the validity of the algorithm output can thus be stated as suitable for the derivation of a long term climatology data set.

3.4 FLS Climatology

The adapted FLS detection scheme was used for processing all MSG data between 2006 and 2015 for each available MSG scene (342'328 in total). In this way, a high resolution data set with a long temporal coverage was produced as a basis for a representative FLS climatology for Europe. The following sections will address spatial patterns, seasonal variations and interannual trends in FLS occurrence.

3.4.1 Spatial Distribution

In Figure 3.6 averaged FLS occurrences derived from all scenes processed in this survey are depicted. The overall average lies at $3.23\,\mathrm{h\,d^{-1}}$ with a maximum reached in Galicia in northwestern Spain ($6.03\,\mathrm{h\,d^{-1}}$) and a minimum in the Sierra Nevada with $0.52\,\mathrm{h\,d^{-1}}$ in southeastern Spain. Total cloud cover measurements by the Spanish Agencia Estatal de Meteorologia show the same spatial pattern for the Iberian peninsula with yearly average minima at the southern and maxima at the northern coast (Sanchez-Lorenzo et al., 2009).

The map shows a general increase towards the Northeast with spatially extensive high values reached in Poland and the Baltic states. Southern regions, especially those close to the Mediterranean coast and mountain ranges like the Alps, the Sistema Central and the Sierra Nevada in Spain as well as the Pyrenees, are characterized by considerably fewer FLS occurrences. The same pattern was already detected by Cermak et al. (2009) for the winter months of the years from 2004 to 2008 and it is confirmed by METAR reports averaged over the period from 2006 to 2015 (Figure 3.7). With an average increase of 0.29 h/d/degree latitude and an R^2 of 0.53 the dependence of FLS occurrence on latitude is clearly reflected in the METAR reports.

Small scale regional variations in FLS frequency are generally closely linked to topographical features. The Po and the Danube valley, the upper Rhine area and the swiss plateau show higher values than their adjacent mountainous areas (Alps, Vosges, Black Forest). These findings are confirmed by regional investigations of Scherrer and Appenzeller (2014) for the swiss plateau and by Fuzzi et al. (1992), Wobrock et al. (1992) as well as Bendix (1994) for the Po valley who also emphasize the topographical dependence of FLS layers.

Local maxima in FLS distribution in regions like Haute-Loire in southern France and southeastern Belgium can also be explained by their topographical characteristics. Here, low lying areas are surrounded by elevated terrain, acting as windshield

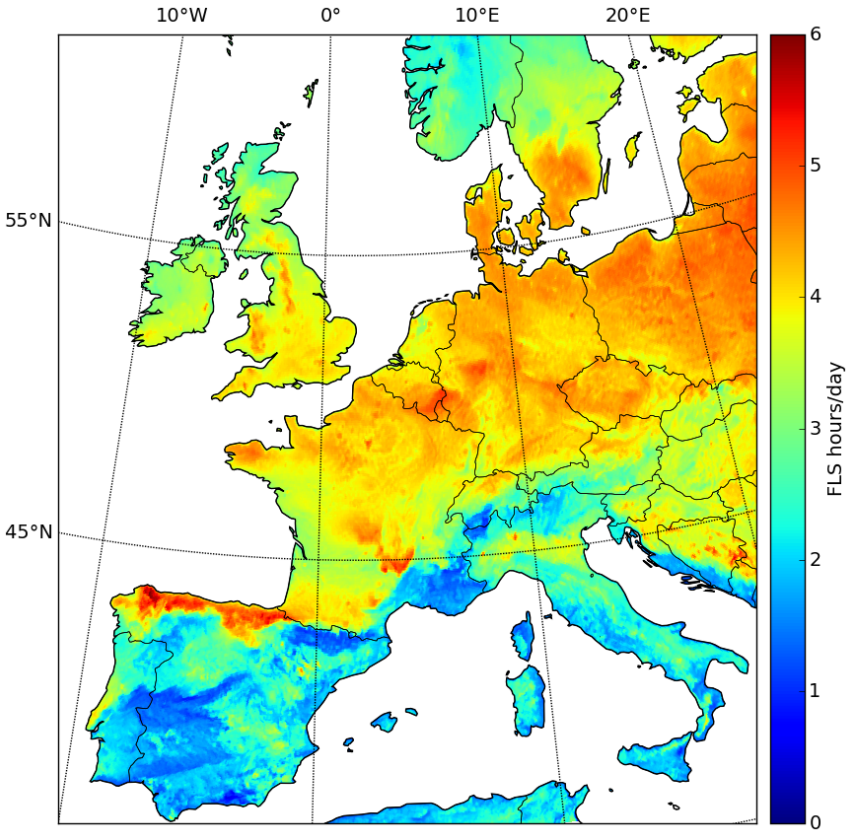

FIGURE 3.6: FLS $\mathrm{h\,d^{-1}}$ averaged from 2006 to 2015 (342'328 scenes)

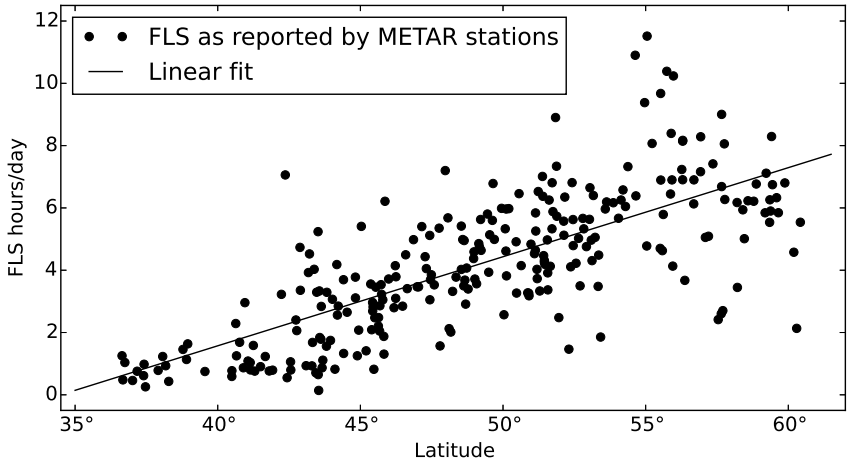

FIGURE 3.7: Latitude-dependence of FLS $\mathrm{h\,d^{-1}}$. Each dot represents one METAR station. Values are averaged over the 10 year period from 2006 to 2015. The slope of the linear fit is calculated as $0.29\,\mathrm{h/d/degree\,latitude}$ with an R^2 of 0.53

and barrier to FLS layers that form in the adjacent depressions due to radiative cooling processes during inverted weather situations and because of confluence of nocturnal katabatic flows with resulting condensation processes. In addition to the Mediterranean coast area, maritime influence also determines local distributions in FLS occurrence in other areas as can be seen in Spain where high FLS frequencies can be found along the entire northern coast of the Iberian Peninsula.

3.4.2 Seasonal Variations

Seasonal variations in FLS occurrence are depicted in Figure 3.8. Averaged values over the complete 10 year period are plotted for (a) winter (Dec-Feb), (b) spring (Mar-May), (c) summer (Jun-Aug) and (d) fall (Sep-Nov). Corresponding total averages of the entire scene are plotted in (e), (f), (g) and (h) separately for each month together with maximum and minimum averages of the respective month during the complete period from 2006 to 2015.

Winter season (Figure 3.8 (a), 3.8 (e)) is characterized by more FLS occurrences than all other seasons which is confirmed by METAR stations reporting the same seasonal variation (cf. Figure 3.4). Average FLS cover prevails on a high level for all months with $4.29\,\mathrm{h\,d^{-1}}$ in December, $4.51\,\mathrm{h\,d^{-1}}$ in January and $4.42\,\mathrm{h\,d^{-1}}$ in February. FLS distinctly decreases from Northeast to Southwest with generally high values in Southern Scandinavia, the Baltic states, Poland, the Czech Republic, parts of Germany and Eastern France. FLS peaks can mostly be found inland where climate is dominated by continental influences and where topographical properties favor calm wind situations as in the Po, Danube and Upper Rhine valley. FLS forms in these regions mainly due to cold air drainage flows from the surrounding terrain, resulting in nightly condensation processes. On the other hand, coastal areas show less FLS occurrences. A possible explanation for this distribution is given by extensive cooling of continental areas in winter with a subsequent emergence of inverted weather situations resulting in radiation fogs at night as well as low stratus layers after lifting of the inversion layer during daytime. This effect is less prominent in coastal regions due to maritime influences preventing the cooling effect to result in condensation of the lower atmospheric layers. Generally low values, on the other hand, are prevalent in Italy, Spain and in the North of Great Britain. Especially mountainous areas like the Alps and the Pyrenees are dominated by very few FLS occurrences. These mountain ranges outreach the vertical extent of FLS layers, thus acting as barriers to the surrounding FLS events, keeping them in the adjacent valleys (cf. Scherrer and Appenzeller, 2014).

In spring (Figure 3.8 (b), 3.8 (f)), FLS occurrences are clearly less frequent than during winter with average values continuously decreasing from $3.46\,\mathrm{h\,d^{-1}}$ in March to $2.96\,\mathrm{h\,d^{-1}}$ in April and to $2.79\,\mathrm{h\,d^{-1}}$ in May. This can mainly be attributed to a decrease of radiation FLS formations as temperatures at night stay above the dew point more often. The overall FLS distribution, however, shows the same characteristics as in winter with smallest values in southern regions as well as mountain ranges and higher values towards the Northeast. One exception to the winterly distribution is visible along the coastal regions of the Atlantic and the North Sea where FLS occurrences are more frequent than in the inland regions. This points to an increase of maritime FLS that gets advected to coastal regions but dissipates inland due to higher temperatures over the continental areas.

During summer (Figure 3.8 (c), 3.8 (g)), FLS decreases to minimal averages of only $2.54\,\mathrm{h\,d^{-1}}$ in June, $2.09\,\mathrm{h\,d^{-1}}$ in July and $2.07\,\mathrm{h\,d^{-1}}$ in August. The overall distribution is very similar to the distribution in spring: Small values in continental areas and high values along the coastlines. Especially the northern coast of Spain shows high

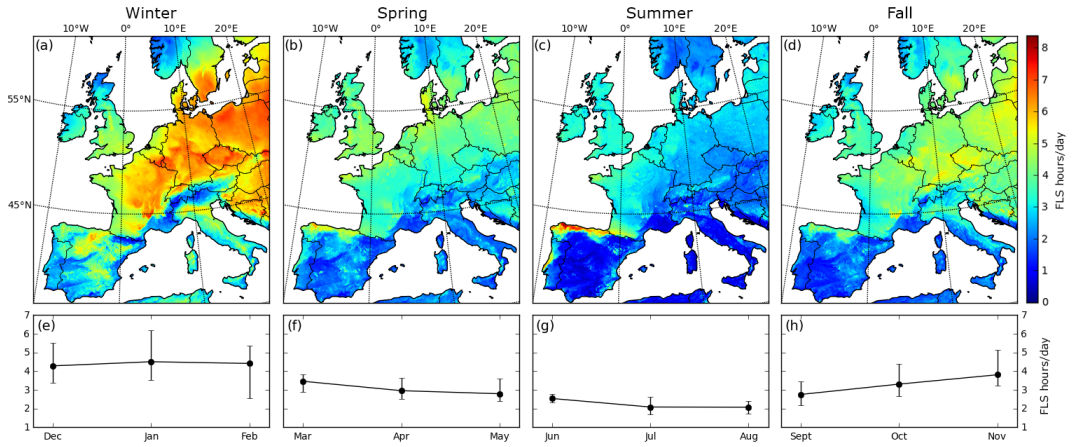

FIGURE 3.8: FLS $\mathrm{h\,d^{-1}}$ averaged over the four seasons. (a) Winter: Dec, Jan, Feb; (b) Spring: Mar, Apr, May; (c) Summer: Jun, Jul, Aug; (d) Fall: Sep, Oct, Nov. Subplots e, f, g and h give corresponding monthly averages with bars denoting maximum and minimum averages of the respective month during the complete period from 2006 to 2015.

FLS frequencies during summer and thus an inverted seasonal trend in comparison to other regions in Europe. A reason for this can be found in a shift of the general wind patterns over the Iberian Peninsula and the Atlantic during summer. Visual analyses of SEVIRI infrared composites show that due to humid and moderately cool temperatures, FLS forms over the Atlantic and is then transported from the Bay of Biscay in Southeast direction towards northern Spain. Once reaching the mainland, the low lying stratus layers are orographically blocked by the Cantabrian mountains and thus are not transported further inland.

After August (Figure 3.8 (d), 3.8 (h)), average FLS values continuously rise again with $2.76\,\mathrm{h\,d^{-1}}$ in September, $3.31\,\mathrm{h\,d^{-1}}$ in October and $3.81\,\mathrm{h\,d^{-1}}$ in November. The frequency distribution in fall is very similar to winter season with a prevalent northwest to southeast decrease and more FLS occurrences towards continental areas, especially in topographically favored areas like river valleys where radiative FLS formations increase again.

3.4.3 Interannual Trends

In order to investigate interannual trends over the 10 year period from 2006 to 2015, monthly averages of FLS frequency were computed for all land pixels in the domain. The results are plotted in Figure 3.9. The plot clearly depicts the seasonal variability in FLS occurrence over the 10 year period. While all other winters show maxima above $5\,\mathrm{h\,d^{-1}}$, the winters of 2011/2012 and 2013/2014 are less strong pronounced. Here, average values do not exceed $4\,\mathrm{h\,d^{-1}}$. The winter of 2011/2012, however, is preceded by a very strong November with $5.13\,\mathrm{h\,d^{-1}}$ which differs by more than two standard deviations from the 10 year average FLS frequency of $3.80\,\mathrm{h\,d^{-1}}$ for November. Reasons for this high FLS occurrence can be found in the general wind and weather patterns over Europe during that time (DWD, 2012a,b). Strong anticyclonic conditions throughout Central Europe and the Baltic states as well as persistent southeastern cold surface winds towards Europe were recorded. These atmospheric conditions caused it to be the driest November in Germany since the beginning of record keeping but also the FLS heaviest November for Europe for the last ten years. Until February 2012 the

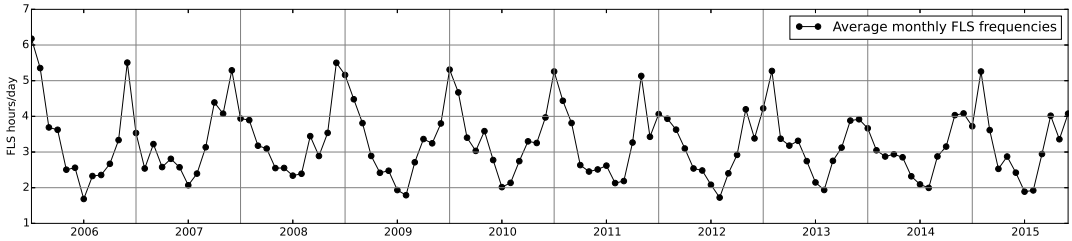

FIGURE 3.9: Average FLS frequency per month for all land pixels in the investigated domain.

synoptic weather situation of the subsequent months was mainly influenced by cyclonic conditions with relatively warm temperatures (DWD, 2013a,b), thus hindering the formation of FLS during the whole winter period of 2011/2012. For the winter of 2013/2014 similar conditions were reported (DWD, 2014, 2015a,b). Fast moving cyclones bringing warm and humid air masses to Europe caused this winter to be influenced by one of the highest precipitation rates since the beginning of the 20th century that consequently came in hand with relatively low FLS frequencies. Besides these peculiarities, a distinct trend is not visible in the time series.

It has, however, to be stated, that a time period of 10 years is insufficient to make reliable statements about long term trends of climatological variables. Consequently, the validity of any future predictions of FLS frequencies in Europe on the basis of this data set is limited.

Possible sub-trends in one or more months of the averaged FLS data (e.g. trends between the Januaries of all 10 years) were tested for by using the seasonal Mann-Kendall trend test. The test was first developed by Hirsch et al. (1982) based on the original test introduced by Mann (1945) and modified by Kendall (1955). It is a nonparametric test to find monotonic trends in a time series that is characterized by seasonal variability and does not require normally distributed data.

First, the Mann-Kendall statistic is calculated separately for each month:

$$S_i = \sum_{k=1}^{n_i-1} \sum_{j=k+1}^{n_i} \mathrm{sgn}(x_{i,j} - x_{i,k}) \tag{3.5}$$

where S_i is the sum of positive differences between all consecutive observations minus the sum of the respective negative differences for month i. n_i is the number of observations for month i and the sgn() function is defined as:

$$\mathrm{sgn}(x_{ij} - x_{ik}) = \begin{cases} 1, & \text{if } x_{ij} - x_{ik} > 0 \\ 0, & \text{if } x_{ij} - x_{ik} = 0 \\ -1 & \text{if } x_{ij} - x_{ik} < 0 \end{cases} \tag{3.6}$$

where x_{ij} and x_{ik} are consecutive observations of month i. Second, the variance for each S_i is calculated via

$$\mathrm{VAR}(S_i) = \frac{1}{18}\left(n_i(n_i-1)(2n_i+5)\right) \tag{3.7}$$

with n_i being the number of observations of month i. The final test statistic is given by

$$Z_{SK} = \begin{cases} \frac{S'-1}{\sqrt{VAR(S')}}, & \text{if } S' > 0 \\ 0, & \text{if } S' = 0 \\ \frac{S'+1}{\sqrt{VAR(S')}}, & \text{if } S' < 0 \end{cases} \tag{3.8}$$

where S′ is defined as

$$S' = \sum_{i=1}^{m} S_i \tag{3.9}$$

and VAR(S′) is defined as

$$VAR(S') = \sum_{i=1}^{m} VAR(S_i) \tag{3.10}$$

where m is the number of all seasons accounted for in the test, which equals to 12 months in this case.

With $|Z_{SK}| = 0.96$ and a Z score of 1.96 for the 97.5 percentile point, H_0 cannot be rejected at a level of significance of 5 %. Thus, a significant monotonic trend cannot be expected in the seasonal time series of the averaged FLS frequencies. This, however, is only valid for the total monthly averages of all land pixels in the investigated domain. To take a closer look at the spatial distribution of possible local trends in FLS frequency, trend maps have been created for each season (Figure 3.10). Trends were computed by fitting a linear function to the average values of the respective season, separately for each pixel in the domain.

$$f = at + b \tag{3.11}$$

where f is average FLS frequency, a is slope, b is intercept and t is time in years. (a), (b), (c) and (d) of Figure 3.10 show a for each pixel in FLS $h\,d^{-1}\,a^{-1}$. Subplots (e), (f), (g) and (h) show the corresponding trends in the METAR records for all 273 stations used in this study. Respective R^2 values of each pixel's linear fit as a measure of reliability of the computed trend are depicted in Figure 3.12 in the online supporting information material. Subplots (i), (j), (k) and (l) depict the evolution of each season's overall average for the investigated 10 year period derived from the FLS product. To further investigate correlations between FLS and its meteorological drivers, interannual air temperature and wind speed variations (2 m above ground) have been visually analyzed accordingly on the basis of all available METAR records between 2006 and 2015. The trends are plotted in Figure 3.11 with (a), (b), (c) and (d) showing temperature trends and (e), (f), (g) and (h) showing the respective wind speed trends.

In winter (Figure 3.10 (a), (e) and (i)), trends in FLS frequencies are stronger than in all other seasons. This can mainly be attributed to the fact that FLS frequencies themselves are higher in winter, resulting in stronger changes of absolute frequency values. Figure 3.10 (a) shows, that vast parts of Central Europe experienced a decrease of FLS occurrences. The northern coastal regions of France, Belgium, the Netherlands, Germany and Poland as well as the swiss plateau and most parts of Great Britain, Italy and southeast Europe are all subject to a considerable decline in FLS frequency with more than $0.25\,h\,d^{-1}\,a^{-1}$ averaged over the investigated 10 year period. Only Scandinavia, Estonia and some minor parts of Portugal and Italy show an increase in FLS occurrence that reach up to $0.20\,h\,d^{-1}\,a^{-1}$ in Sweden. Noticeably,

decreasing trends generally go in hand with high R^2 values (often more than 0.5) whereas increasing trends are much less reliable with R^2 values in Scandinavia and the Baltic states not exceeding 0.2 in most parts (Figure F1a).

Compared to the FLS product, trends derived from METAR reports show a more heterogeneous picture but generally the same pattern (Figure 3.10 (e)). Most stations reported a negative trend in winter. Especially in Great Britain, the Atlantic coast of France, the Netherlands and Germany as well as the Adriatic coast and continental areas in Spain and Eastern Europe show very similar distributions to the satellite derived product. Also, regional positive trends like in the western parts of the Iberian Peninsula and in Scandinavia are resembled in the METAR reports. On the other hand, some local frequency increases in northern France and in the Balkans, that are visible in the METAR data, are not grasped by the FLS product. These discrepancies are presumably a result of the different spatial features of both datasets. METAR data represent local properties only whereas the coarse resolution of the SEVIRI instrument leads to a smoother frequency distribution in the FLS product. Especially in complex terrain, this can lead to local differences in both datasets.

Subplot 3.10 (i) confirms the general decrease in FLS frequency in winter seasons over the investigated period with a maximum average FLS occurrence of $5.77\,\mathrm{h\,d^{-1}}$ in 2006 and a minimum of $3.55\,\mathrm{h\,d^{-1}}$ in 2014. Although a distinction between different FLS types is not possible on the basis of this study's product, it can be assumed that the general decrease of FLS can be ascribed to a decline of predominantly in winter occurring stationary FLS events induced by radiative cooling. This assumption is also backed up by trends in temperature and wind speed (Figure 3.11 (a) and (e)). Temperature values generally increased with up to more than $0.25\,^{\circ}\mathrm{C/a}$ in Eastern Europe. Concurrently, wind speeds also increased, especially over Great Britain. Assuming otherwise unchanged conditions, both, increases in temperature and wind speed, lead to a decrease of radiative FLS events as they are bound to calm atmospheric conditions and a distinct nightly cooling of the boundary layer below the dew point. With higher temperatures and higher wind speeds, consequently these conditions emerge less frequent.

In spring, no long term trend can be detected (Figure 3.10 (j)). Total averages fluctuate closely around $3.00\,\mathrm{h\,d^{-1}}$ over the complete 10 year period. However, slight spatial patterns are visible in the FLS product (Figure 3.10 (b)) with strongest inclines in Great Britain, the Benelux countries and around Southeast Germany, the Czech Republic as well as northern Austria. Declining frequencies can be found in Scandinavia and the Baltic states. The general trend patterns can also be seen in the METAR data showing a similar distribution for all these regions with inclines in Great Britain and Central Europe as well as declines in Scandinavia and the Baltic states (Figure 3.10 (f)). While coefficients of determination (Figure F1b) are high, both for decreasing and increasing trends, regions with no distinct trend show low R^2 values. In these regions, the variation from year to year superimposes a possible general pattern in the data. Temperature and wind speed trend distributions closely follow the FLS trend distribution (Figure 3.11 (b) and (f)). Increases in temperature and wind speed in Scandinavia and Eastern Europe go in hand with decreases in FLS occurrence. On the other hand, decreases in temperature and wind speed in Great Britain and Central Europe coincide with increasing FLS frequencies, again pointing to a close dependence of FLS occurrence on these meteorological conditions.

Trends in summer are similarly weak as in spring with FLS averages remaining on approximately the same level of $2.23\,\mathrm{h\,d^{-1}}$ throughout the whole 10 year period (Figure 3.10 (k)). The spatial pattern, however, is almost inverted to spring season (Figure 3.10 (c))). Increasing trends in summer mainly occur in regions with

decreasing trends in spring and vice versa (e.g. Portugal, France, Great Britain, Germany, Austria, Scandinavia). Especially in France, decreasing trends go in hand with high R^2 values, indicating the reliability of the computed decrease in this region (Figure 3.12 (c)). METAR reports (Figure 3.10 (g)) show the same trend inversion for Great Britain, Scandinavia and Germany but are generally more heterogeneously distributed during the summer season with weak increasing an decreasing trends in adjacent regions. As in summer FLS events are less frequent, than in all other seasons and because most FLS events can be ascribed to advective rather than radiative FLS situations, temperature and wind speed trend patterns show weaker correlations to FLS trend patterns in this season (Figure 3.11 (c) and (g)). However, temperature increases in Central Europe still go in hand with FLS frequency decreases in these regions, whereas decreasing temperatures coincide with FLS frequency increases in Scandinavia.

In fall, differences in trends are more pronounced than in spring and in summer. While there is a slight overall incline of FLS occurrences in fall (Figure 3.10 (l)), this can mainly be ascribed to a small region in Central Europe around southern Germany, the Czech Republic and the Balkan states (Figure 3.10 (d)). Most other parts are characterized by very little changes and small R^2 values (Figure F1 (d)). The weak positive trends in Scandinavia, the Benelux countries and Great Britain that are depicted in the data derived from the FLS product are more pronounced in the METAR data where they reach values of more than $0.25\,\mathrm{h\,d^{-1}\,a^{-1}}$ for the vast majority of stations (Figure 3.10 (h)). In general, however, both datasets show similar trend distributions for most regions. While wind speeds slightly decreased by $0.03\,\mathrm{m\,s^{-1}\,a^{-1}}$ (Figure 3.11 (d) and (h)), there are no distinct regional patterns in this variable. On the other hand, the temperature trend distribution shows obvious patterns with increases in South Europe (Spain, southern France, Adriatic coast) and slight decreases in Great Britain and the Benelux countries, thus closely resembling the FLS distribution with decreasing FLS frequencies in South Europe and increasing FLS frequencies in the Benelux and Great Britain.

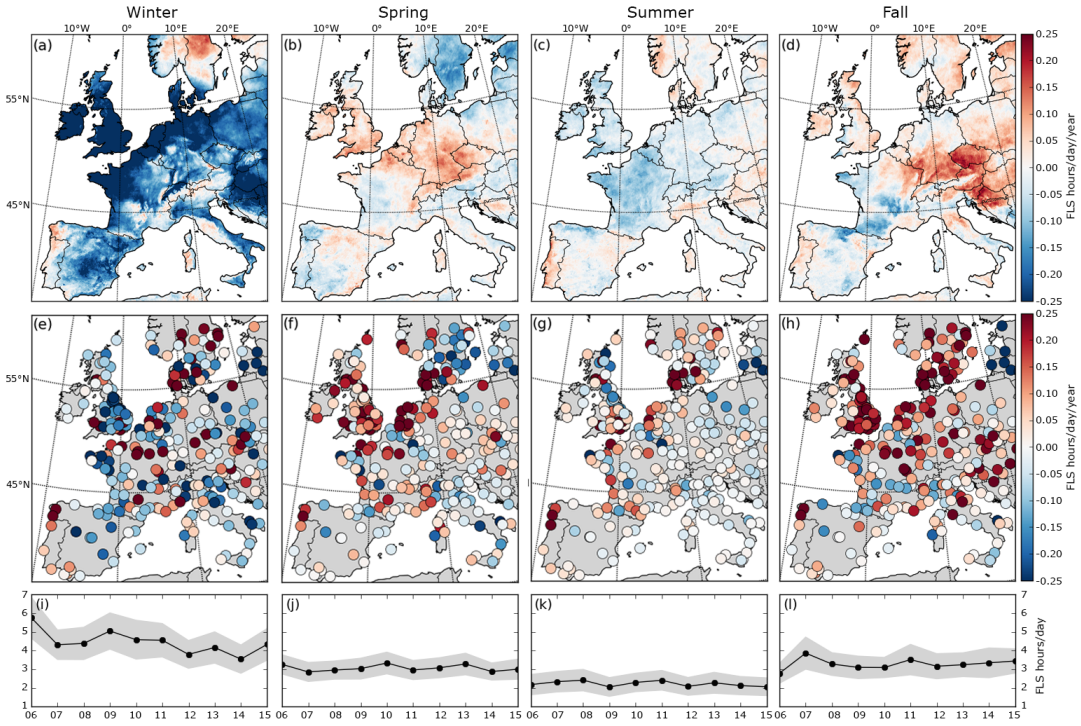

FIGURE 3.10: Trends in FLS $\mathrm{h\,d^{-1}}$ for the period between 2006 and 2015 averaged over the four seasons. Trends are given as the gradient of a linear regression fitted to each season's average. a, b, c and d show changes in FLS $\mathrm{h\,d^{-1}}$ over one year derived from the generated FLS product. e, f, g and h show the corresponding trend values of all 273 METAR stations. i, j, k and l show overall season averages (black line) together with standard deviations (gray area).

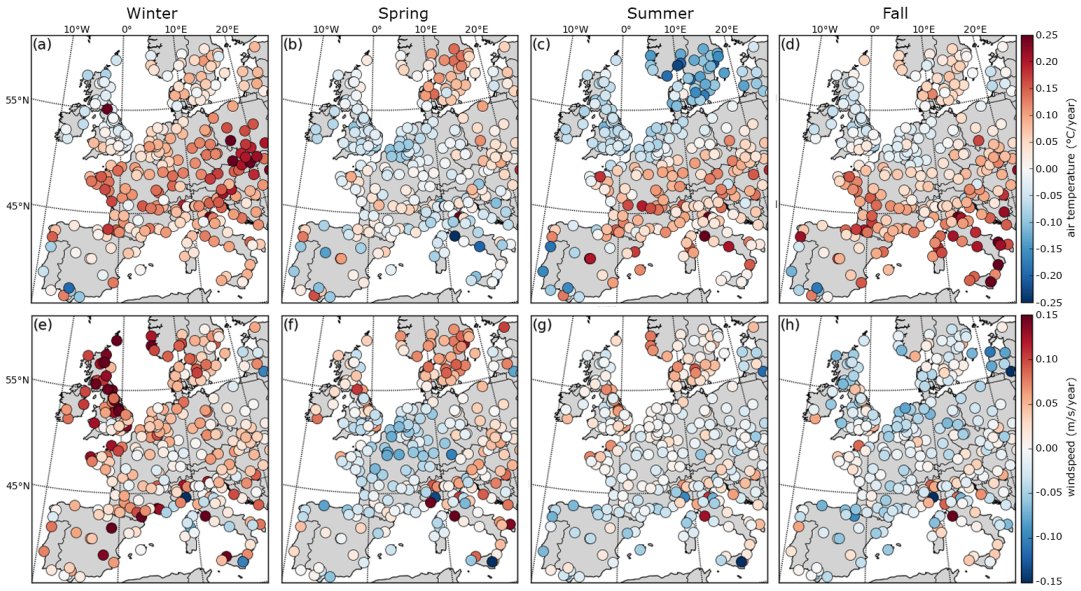

FIGURE 3.11: Trends in 2 m air temperature and windspeed for the period between 2006 and 2015 averaged over the four seasons. Trends are given as the gradient of a linear regression fitted to each season's average. a, b, c and d show changes in temperature (°C) over one year derived from METAR data. Negative values signify a decrease, positive values signify an increase. e, f, g and h show the corresponding wind speed trends ($\mathrm{m\,s^{-1}}$).

3.5 Conclusion

The data set presented in this study provides information about FLS frequency distributions in Europe on a level of detail that to our knowledge hasn't been reached before. Spatially explicit data of FLS occurrence is given at a temporal resolution of 15 min for the complete 10 year period reaching from 2006 to 2015. A comparison to ground truth data (METAR reports) shows reasonable agreement with slight variations over the annual cycle.

Spatial average patterns in FLS occurrence confirm recent studies (Cermak et al., 2009) and show a distinct Northeast-Southwest gradient as well as topographically contingent high- and low frequency regions.

Seasonal average maps show a strong dependence of FLS frequencies and their spatial distribution on the annual cycle. While winter months are characterized by high values and a pronounced latitudinal gradient, in summer frequencies are substantially lower with Atlantic coast regions gaining in importance. Recent studies on annual variations in fog, haze and low stratus occurrence show comparable results and find highest frequencies mainly during anticyclonic weather situations in cold and dry seasons, especially over continental areas (Belorid et al., 2014; Chen and Wang, 2015).

Interannual trends are most pronounced in winter with vast areas of Central Europe showing a distinct decrease in FLS occurrence over the 10 year period. A visual comparison to trends in the meteorological drivers of FLS showed a close dependence of FLS on temperature and wind speed, especially in winter when radiative FLS events prevail. These findings are in good agreement with latest investigations of long term trends in fog, mist and haze frequencies, both in Europe (Vautard et al., 2009) and worldwide (Belorid et al., 2014; Klemm and Lin, 2016).

One reason for the general decline of fog and low stratus in winter may be given by rising temperatures in the lower atmosphere due to climatological changes (Vautard et al., 2009). At higher temperatures but otherwise unchanged conditions, dew points are reached less often, causing less condensation of water vapor and thus less FLS formation. Also, air quality improvements lead to a reduction of aerosols that act as condensation nuclei in supersaturated conditions. This amplifies the reduction in FLS formation additionally (Klemm and Lin, 2016; Vautard et al., 2009).

The FLS data set produced in this study will contribute to a deeper understanding of FLS frequency distributions and atmospheric processes coupled to FLS formation and dissipation. As mentioned before, a time period of 10 years is, however, not sufficient to make reliable statements about future developments of FLS frequency distributions. Climate research is dependent on harmonious data series over long time periods. A future extension of the FLS data set in combination with ground based data records reaching further into the past will thus provide valuable information for applications in this field.

As meteorological services depend on base height information of detected FLS layers to improve short term forecasts of fog events, the FLS product will be used as a first step towards a ground fog product in future applications. This has already been tested successfully for daytime scenes in Cermak and Bendix (2011) and, with some adaptation, could be applied to a long term data set similar to this study. Currently, the feasibility of a hybrid approach including physical threshold tests and machine learning models using ground truth records as training data to generate a ground fog product is investigated in our working group. If successful, by this means, forecasting products could be validated spatially continuously against a long term ground fog data set to achieve better results in prediction outputs.

Acknowledgements

This study was performed within the project "Ground fog detection and analysis with Machine Learning" (GFog-ML), generously funded by the German Research Foundation (DFG). The authors also thank EUMETSAT for providing the MSG satellite data used in this study.

Supporting information

Year	Month	a	b	c	d
2006	1	100201	29557	67308	298230
	2	90003	21827	67171	289141
	3	48668	17831	57907	377107
	4	33774	22992	31043	427557
	5	21664	14634	29756	470305
	6	17384	23014	16230	528178
	7	13494	16469	14277	611568
	8	18634	13551	29654	444442
	9	24105	14291	31995	469386
	10	39086	13677	51117	438449
	11	43011	13471	60401	404044
	12	102466	24658	71440	318825
2007	1	43793	16945	76594	350291
	2	4047	1371	16613	49032
	3	36275	17552	38724	421132
	4	22509	19702	20809	556409
	5	35176	15523	35064	444012
	6	29982	15193	29119	431288
	7	15377	13044	27740	492680
	8	23499	15270	29060	500009
	9	25260	22318	30387	456106
	10	64384	28366	54004	420076
	11	54473	21832	64543	369668
	12	114537	18959	81481	347131
2008	1	63855	17310	95718	343152
	2	59600	17347	58785	414970
	3	24871	16048	49288	392915
	4	27833	16952	36880	434067
	5	13331	16695	17342	440826
	6	15192	19640	18230	489450
	7	17278	18863	19143	523944
	8	20957	15910	31739	478383
	9	23603	29445	26868	444033
	10	30126	15297	49896	435920
	11	52213	11801	79919	349448
	12	106325	18770	81066	303879
2009	1	103784	20312	88580	345887
	2	60293	18687	75383	328435
	3	46048	19986	54936	432913
	4	31453	18001	28921	506931
	5	21733	18237	26223	543295
	6	21015	17566	22992	505321
	7	11779	13732	18954	550196
	8	15921	13445	22193	581795
	9	29245	20292	30309	514110
	10	37329	20750	58534	453928
	11	50610	10892	79775	355512
	12	59083	16925	90089	351700
2010	1	100898	20300	98173	310271
	2	72689	20710	76409	291304
	3	38908	24561	52134	463609
	4	28753	20567	22694	528570
	5	30899	24500	35591	446187
	6	22593	20818	27255	512544
	7	13972	18240	24907	565946
	8	16766	14112	32802	508626
	9	22523	16549	39319	480654

Continued on next page

Table 3.1 – *Continued from previous page*

Year	Month	a	b	c	d
	10	46935	19550	47997	475891
	11	47249	11769	83690	336977
	12	76403	13423	103218	344585
2011	1	95455	25697	108201	339667
	2	82070	16896	86876	346660
	3	60405	24490	45797	489317
	4	24418	21381	33221	575290
	5	16200	18271	19169	578457
	6	18895	18389	26542	498347
	7	25222	17157	33708	494917
	8	16588	16549	27929	540539
	9	16694	15887	33244	531824
	10	46249	19247	52129	494350
	11	130271	14237	86187	383957
	12	45865	12457	85683	381155
2012	1	57064	24678	87173	402339
	2	56340	24701	73956	418382
	3	72139	19660	42295	547247
	4	26909	16469	41030	426917
	5	28720	19064	32864	542063
	6	24363	16485	36013	478625
	7	16165	15854	31049	523352
	8	13420	14013	23157	591658
	9	19862	17529	31718	514659
	10	43440	13142	67243	445456
	11	68041	18409	83730	364158
	12	51188	12897	103938	361196
2013	1	79731	19392	132297	317249
	2	58004	19181	54851	225821
	3	45071	25500	63309	407837
	4	32554	23414	40298	491374
	5	33934	18754	38686	452905
	6	33376	17620	33243	500162
	7	22053	13851	17291	459310
	8	15221	15991	21361	594952
	9	30614	16201	40808	483413
	10	38091	22473	64358	441913
	11	47988	19185	79693	371204
	12	75525	17077	86048	432286
2014	1	50746	10084	82150	264574
	2	37488	12032	70239	364689
	3	42898	16076	49255	533615
	4	29402	21073	42520	502763
	5	26244	20400	33402	496987
	6	18243	18900	18344	548969
	7	20133	13605	27115	550363
	8	11130	11656	22958	535151
	9	35568	25841	36581	512522
	10	44515	17255	68078	455825
	11	73313	13911	104290	347111
	12	48183	14764	79803	310553
2015	1	60065	15183	101335	377132
	2	94498	18838	61035	345569
	3	50355	23753	57041	471100
	4	23573	16287	29206	520627
	5	25775	22050	34344	500395
	6	23724	19447	22631	541147
	7	15354	16640	26931	572026
	8	12994	18302	24466	591291
	9	22014	22926	25765	505106
	10	54644	28954	64562	440786
	11	52075	11559	80973	355160
	12	65001	19612	82272	329954

TABLE 3.1: Validation results. Values of a, b, c and d are given for each month between 2006 and 2015.

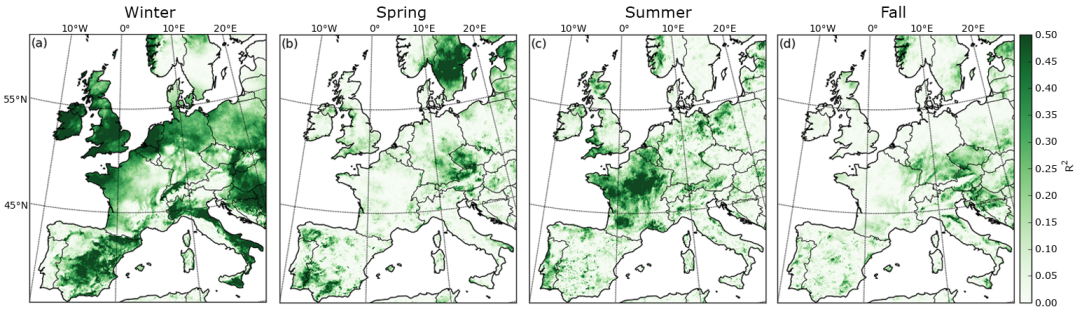

FIGURE 3.12: R^2 values corresponding to the linear fits of FLS frequencies depicted in Figure 3.10 (a-d).

Bibliography

Aminou, D. (2002). MSG 's SEVIRI Instrument. *ESA Bulletin*, 111:15–17.

Aminou, D. M. A., Jacquet, B., and Pasternak, F. (1997). Characteristics of the Meteosat Second Generation (MSG) radiometer/imager: SEVIRI. In Fujisada, H., editor, *SPIE 3221, Sensors, Systems, and Next-Generation Satellites*, pages 19–31, London, United Kingdom.

Belorid, M., Lee, C. B., Kim, J. C., and Cheon, T. H. (2014). Distribution and long-term trends in various fog types over South Korea. *Theoretical and Applied Climatology*, 122(3-4):699–710.

Bendix, J. (1994). Fog climatology of the Po Valley. *Rivista di Meteorologia Aeronautica*, 54(3-4):25–36.

Bendix, J. (2002). Working Group I , National Report Germany. Technical report, COST 722.

Bendix, J. and Bachmann, M. (1991). Ein operationell einsetzbares Verfahren zur Nebelerkennung auf der Basis von AVHRR-Daten der NOAA-Satelliten. *Meteorologische Rundschau*, 43:169–178.

Bendix, J., Eugster, W., and Klemm, O. (2011). Fog – boon or bane? *Erdkunde*, 65(3):229–232.

Best, A. C. (1951). Drop-size distribution in cloud and fog. *Quarterly Journal of the Royal Meteorological Society*, 77(3):418–426.

Bruijnzeel, S., Eugster, W., and Burkard, R. (2005). Fog as a Hydrologic Input. In *Encyclopedia of Hydrological Sciences*, pages 559–582. John Wiley & Sons, Ltd, Chichester, UK.

Cermak, J. (2006). *SOFOS - A new Satellite-based Operational Fog Observation Scheme*. PhD thesis, Philipps-University of Marburg.

Cermak, J. and Bendix, J. (2007). Dynamical nighttime fog/low stratus detection based on Meteosat SEVIRI data: A feasibility study. *Pure and Applied Geophysics*, 164(6-7):1179–1192.

Cermak, J. and Bendix, J. (2008). A novel approach to fog/low stratus detection using Meteosat 8 data. *Atmospheric Research*, 87(3-4):279–292.

Cermak, J. and Bendix, J. (2011). Detecting ground fog from space – a microphysics-based approach. *International Journal of Remote Sensing*, 32(12):3345–3371.

Cermak, J., Eastman, R. M., Bendix, J., and Warren, S. G. (2009). European climatology of fog and low stratus based on geostationary satellite observations. *Quarterly Journal of the Royal Meteorological Society*, 135(645):2125–2130.

Chen, H. and Wang, H. (2015). Haze Days in North China and the associated atmospheric circulations based on daily visibility data from 1960 to 2012. *Journal of Geophysical Research: Atmospheres*, 120(12):5895–5909.

Choularton, T. W., Fullarton, G., Latham, J., Mill, C. S., Smith, M. H., and Stromberg, I. M. (1981). A field study of radiation fog in meppen, West Germany. *Quarterly Journal of the Royal Meteorological Society*, 107(452):381–394.

Dozier, J. (1989). Spectral signature of alpine snow cover from the landsat thematic mapper. *Remote Sensing of Environment*, 28(C):9–22.

DWD (2012a). Großwetterlage (Januar bis Dezember 2011). Technical report, Deutscher Wetterdienst, Offenbach.

DWD (2012b). Klimastatusbericht 2011. Technical report, Deutscher Wetterdienst, Offenbach.

DWD (2013a). Großwetterlage (Januar bis Dezember 2012). Technical report, Deutscher Wetterdienst, Offenbach.

DWD (2013b). Klimastatusbericht 2012. Technical report, Deutscher Wetterdienst, Offenbach.

DWD (2014). Großwetterlage (Januar bis Dezember 2013). Technical report, Deutscher Wetterdienst, Offenbach.

DWD (2015a). Großwetterlage (Januar bis Dezember 2014). Technical report, Deutscher Wetterdienst, Offenbach.

DWD (2015b). Klimastatusbericht 2014. Technical report, Deutscher Wetterdienst, Offenbach.

Egli, S., Maier, F., Bendix, J., and Thies, B. (2015). Vertical distribution of microphysical properties in radiation fogs — A case study. *Atmospheric Research*, 151:130–145.

EUMETSAT (2013). MSG Level 1.5 Image Data Format Description. Technical report, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt.

Fuzzi, S., Facchini, M. C., Orsi, G., Lind, J. A., Wobrock, W., Kessel, M., Maser, R., Jaeschke, W., Enderle, K. H., Arends, B. G., Berner, A., Solly, I., Kruisz, C., Reischl, G., Pahl, S., Kaminski, U., Winkler, P., Ogren, J. A., Noone, K. J., Hallberg, A., Fierlinger-oberlinninger, H., Puxbaum, H., Marzorati, A., Hansson, H.-C., Wiedensohler, A., Svenningsson, I. B., Martinsson, B. G., Schell, D., and Georgii, H. W. (1992). The Po Valley Fog Experiment 1989. *Tellus B*, 44(5):448–468.

Giulianelli, L., Gilardoni, S., Tarozzi, L., Rinaldi, M., Decesari, S., Carbone, C., Facchini, M., and Fuzzi, S. (2014). Fog occurrence and chemical composition in the Po valley over the last twenty years. *Atmospheric Environment*, 98:394–401.

Glickman, T. S. (2000). *Glossary of Meteorology (2nd ed.)*. American Meteorological Society, Boston, 2 edition.

Güls, I. and Bendix, J. (1996). Fog detection and fog mapping using low cost Meteosat-WEFAX transmission. *Meteorological Applications*, 3(2):179–187.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978.

Hirsch, R. M., Slack, J. R., and Smith, R. A. (1982). Techniques of trend analysis for monthly water quality data. *Water Resources Research*, 18(1):107–121.

Hunt, G. E. (1973). Radiative properties of terrestrial clouds at visible and infrared thermal window wavelengths. *Quarterly Journal of the Royal Meteorological Society*, 99(420):346–369.

Kendall, M. G. (1955). *Rank correlation methods*, volume 3. Charles Griffin, London, 4 edition.

Klemm, O. and Lin, N. (2016). What Causes Observed Fog Trends: Air Quality or Climate Change? *Aerosol and Air Quality Research*, 16(5):1131–1142.

Mann, H. B. (1945). Nonparametric Tests Against Trend. *Econometrica*, 13(3):245–259.

Musial, J., Hüsler, F., Sütterlin, M., Neuhaus, C., and Wunderle, S. (2014). Daytime Low Stratiform Cloud Detection on AVHRR Imagery. *Remote Sensing*, 6(6):5124–5150.

Nemery, B., Hoet, P. H. M., and Nemmar, A. (2001). The Meuse Valley fog of 1930: An air pollution disaster. *Lancet*, 357(9257):704–708.

Obregon, A., Gehrig-Downie, C., Gradstein, S. R., and Bendix, J. (2014). The potential distribution of tropical lowland cloud forest as revealed by a novel MODIS-based fog/low stratus night-time detection scheme. *Remote Sensing of Environment*, 155:312–324.

Pilié, R. J., Mack, E. J., Kocmond, W. C., Rogers, C. W., and Eadie, W. J. (1975). The Life Cycle of Valley Fog. Part I: Micrometeorological Characteristics. *Journal of Applied Meteorology*, 14(3):347–363.

Roach, W. (1995). Back to Basics: Fog: Part 2 - The formation and dissipation of land fog. *Weather*, 50(1):7–11.

Robert A. Houze, J. (1993). *Cloud dynamics*. Academic Press, Inc., San Diego, California.

Sachweh, M. and Koepke, P. (1995). Radiation fog and urban climate. *Geophysical Research Letters*, 22(9):1073–1076.

Sachweh, M. and Koepke, P. (1997). Fog dynamics in an urbanized area. *Theoretical and Applied Climatology*, 58(1-2):87–93.

Sanchez-Lorenzo, A., Calbo, J., Brunetti, M., and Deser, C. (2009). Dimming/brightening over the Iberian Peninsula: Trends in sunshine duration and cloud cover and their relations with atmospheric circulation. *Journal of Geophysical Research Atmospheres*, 114(8):D00D09.

Saunders, R. W. and Kriebel, K. T. (1988). An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing*, 9(1):123–150.

Scherrer, S. C. and Appenzeller, C. (2014). Fog and low stratus over the Swiss Plateau-a climatological study. *International Journal of Climatology*, 34(3):678–686.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A. (2002). An Introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, 83(7):977–992.

Schulze-Neuhoff, H. (1976). Nebelfeinanalyse mittels zusätzlicher 420 Klimastationen - Taktische Analyse 1:2 statt 1:5 Mill. *Meteorologische Rundschau*, 29(3):75–84.

Stone, J. E., Gohara, D., and Shi, G. (2010). OpenCL: A parallel programming standard for heterogeneous computing systems. *Computing in Science and Engineering*, 12(3):66–72.

Vautard, R., Yiou, P., and van Oldenborgh, G. J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nature Geoscience*, 2(2):115–119.

Von Glasow, R. and Bott, A. (1999). Interaction of radiation fog with tall vegetation. *Atmospheric Environment*, 33(9):1333–1346.

Wobrock, W., Schell, D., Maser, R., Kessel, M., Jaeschke, W., Fuzzi, S., Facchini, M. C., Orsi, G., Marzorati, a., Winkler, P., Arends, B. G., and Bendix, J. (1992). Meteorological characteristics of the Po Valley fog. *Tellus B*, 44(5):469–488.

Chapter 4

A Hybrid Approach for Fog Retrieval Based on a Combination of Satellite and Ground Truth Data

This chapter has been published as: Egli, S., Thies, B., & Bendix, J. (2018). A Hybrid Approach for Fog Retrieval Based on a Combination of Satellite and Ground Truth Data. Remote Sensing, 10(4), 26.

Received: 09 February 2018 / Accepted: 16 April 2018
https://doi.org/10.3390/rs10040628

Abstract

Fog has a substantial influence on various ecosystems and it impacts economy, traffic systems and human life in many ways. In order to be able to deal with the large number of influence factors, a spatially explicit high-resoluted data set of fog frequency distribution is needed. In this study, a hybrid approach for fog retrieval based on Meteosat Second Generation (MSG) data and ground truth data is presented. The method is based on a random forest (RF) machine learning model that is trained with cloud base altitude (CBA) observations from Meteorological Aviation Routine Weather Reports (METAR) as well as synoptic weather observations (SYNOP). Fog is assumed where the model predicts CBA values below a dynamically derived threshold above the terrain elevation. Cross validation results show good accordance with observation data with a mean absolute error of 298 m in CBA values and an average Heidke Skill Score of 0.58 for fog occurrence. Using this technique, a 10 year fog baseline climatology with a temporal resolution of 15 min was derived for Europe for the period from 2006 to 2015. Spatial and temporal variations in fog frequency are analyzed. Highest average fog occurrences are observed in mountainous regions with maxima in spring and summer. Plains and lowlands show less overall fog occurrence but strong positive anomalies in autumn and winter.

4.1 Introduction

Fog influences human life in various ways. Due to visibility reduction it can have a hindering and sometimes even lethal impact on air, sea and road traffic (Bendix et al., 2011). As fog also influences the radiation balance in the atmospheric boundary layer by reflecting the incoming solar radiation it often leads to persistent atmospheric

inversions. This is especially true for the winter season in Europe when stable boundary layer conditions occur more frequently. As air pollutants can be trapped over long time periods below the inversion layer the persistence of these conditions can lead to severe air quality reductions that are sometimes even hazardous in industrial agglomerations (Nemery et al., 2001). Power grid stability can also be affected by fog and low stratus (LST) as photovoltaic power generation strongly depends on the cloudiness level. Fast changing fog and LST conditions are hardly predictable and frequently pose a challenge for maintaining grid stability (Köhler et al., 2017). All these factors lead to human casualties and financial losses that range on the same level as those linked to extreme weather conditions, such as tornadoes and hurricanes (Gultepe et al., 2007).

Besides these negative effects, fog plays a vital role in various ecological systems. In areas like the coast redwood region of California, along the coasts of the Atacama Desert in Northern Chile and in the Namib Desert, it supplies otherwise arid ecosystems with life sustaining moisture (Henschel and Seely, 2008; Johnstone and Dawson, 2010; Lehnert et al., 2018; Pinto et al., 2001). In the context of climate change, fog frequency variations could also affect the long term development of regional atmospheric conditions. For instance, Vautard et al. (2009) estimate that about 50% of the warming in Eastern Europe can be attributed to the reduction of low visibility conditions.

Due to the stated risks and potentials resulting from fog conditions there is a growing demand for an extensive data set of fog frequency distributions. A long-term spatiotemporally high-resoluted and spatially explicit fog climatology could significantly improve global climate model simulations and regional fog risk estimations (Bendix, 2002; Duda et al., 1996).

Unfortunately, no such data set exists. The only long term fog climatologies that have been developed recently were simply derived using interpolation techniques on observational in situ measurements for single countries (García-García and Zarraluqui, 2008) or the climatology was limited to single locations (Scherrer and Appenzeller, 2014; Tardif and Rasmussen, 2007). Avotniece et al. (2015) used a combination of in situ measurements and a geostationary satellite cloud product to derive fog frequencies. However, only seasonal average cloud products were used and the computed output was limited to Latvia. In general it can be stated that operational observation networks are too sparsely spread to be able to accurately cover the space between the ground truth stations with reliable fog frequency estimations alone.

Remote sensing data of the Moderate Resolution Imaging Spectroradiometer (MODIS) or of the Advanced Very High Resolution Radiometer (AVHRR) could help fill these gaps with a horizontal resolution of 1 km (Bendix and Bachmann, 1991; Bendix, 2002; Musial et al., 2014). However, as these instruments are mounted on polar-orbiting satellite systems, the temporal resolution is severely limited. Geostationary satellite systems, albeit equipped with less spatial resolution, are more suitable to provide long term continuous data in short measurement intervals. Data records of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) aboard the Meteosat Second Generation (MSG) satellites reach back to 29 January 2004 and thus offer the longest continuous high frequent spatially explicit meteorological satellite data set for Europe.

Cermak and Bendix (2008) developed a fog and low stratus (FLS) detection scheme on the basis of SEVIRI data that was mainly based on a sequence of threshold tests on different band combinations. They applied the scheme to the winter months (December–February) between 2004 and 2008 (Cermak et al., 2009). The original algorithms were later adapted by Egli et al. (2017) to make them suitable for varying

conditions in different domain regions and for seasonal changes. Unfortunately, all satellite-based approaches suffer from the same insufficiency as they can only provide information from the top perspective. The approach of Cermak and Bendix (2008) and Egli et al. (2017) thus did not provide data that distinguished between fog and elevated low stratus layers.

Fog, however, as internationally defined, is a ground touching cloud that reduces horizontal visibility (VIS) to less than 1 km (Glickman, 2000). In order to be able to derive fog information from satellite data Cermak and Bendix (2011) developed a microphysics-based approach that was tuned for radiation fog conditions. Estimated cloud liquid water paths and cloud top altitudes are used as input to the procedure to iteratively calculate cloud base altitudes (CBA). These CBAs are then merged with a digital elevation model (DEM) to derive areas of ground touching clouds (=fog). Although this method showed satisfying validation results for a selected sample of MSG records in the autumn months of 2005, a general application on a multiannual data set was not performed. This was due to the fact that the method is only applicable to radiation fog conditions during daytime with sun elevations above 10°. The computed data set would thus be temporally and spatially discontinuous, especially in winter months when sun elevations are too low throughout the complete diurnal cycle in higher latitudes. Also, Egli et al. (2015) showed that the microphysical assumptions of this approach are not always met during different life cycle stages of radiation fogs which causes large errors in the CBA retrieval. As the methods of Kawamoto and Nakajima (2001) that are used for LWP derivation are computationally very expensive, the application on a large data set is additionally impeded.

Remote sensing fog detection algorithms that are based on physical assumptions alone thus have difficulties to adequately describe the dynamical interactions of all influence factors that cause fog formation and dissipation. Machine learning (ML) approaches are better suited to exploit the complete information content of the input variables provided. To circumvent the insufficiencies of previous fog detection attempts, a novel satellite and ground truth data hybrid approach for fog detection was developed and implemented in this study. In principle, for each MSG scene a machine learning model is trained using SEVIRI data and terrain information as independent inputs. CBA observations of Meteorological Aviation Routine Weather Reports (METAR) and synoptic weather observations (SYNOP) from the German Weather Service (DWD) were specified as target feature. The trained models were then used to predict CBA values for all cloud-covered regions in the complete domain. By applying appropriate thresholds to the derived CBA layers, regions with fog occurrence were identified.

The procedure was applied to MSG data covering Europe (WMO region VI) from 2006 to 2015 to generate a continuous baseline fog climatology with a high spatiotemporal resolution. Using the hybrid approach, significantly less computational power is needed in comparison to the microphysics based approaches which speeds up the processing time by a factor of about 100. Also, the procedure is not limited to radiation fogs but is generally applicable to all fog conditions, e.g. advection, orographic or coastal fogs. The main benefit, however, is the general applicability of the procedure to all sun elevations even below 10°. This becomes evident by looking at the diurnal and annual frequency changes in the fog station data (see Figure 4.1). Both, hourly and monthly fog frequency averages show a distinct pattern. Due to the usual temperature decrease after sunset and generally low temperatures during the winter months, condensation on average sets in at lower levels which in turn increases the probability of fog occurrence. As this leads to a strong fog increase during night times and in the winter months the usability of the approach at low sun elevations even beyond

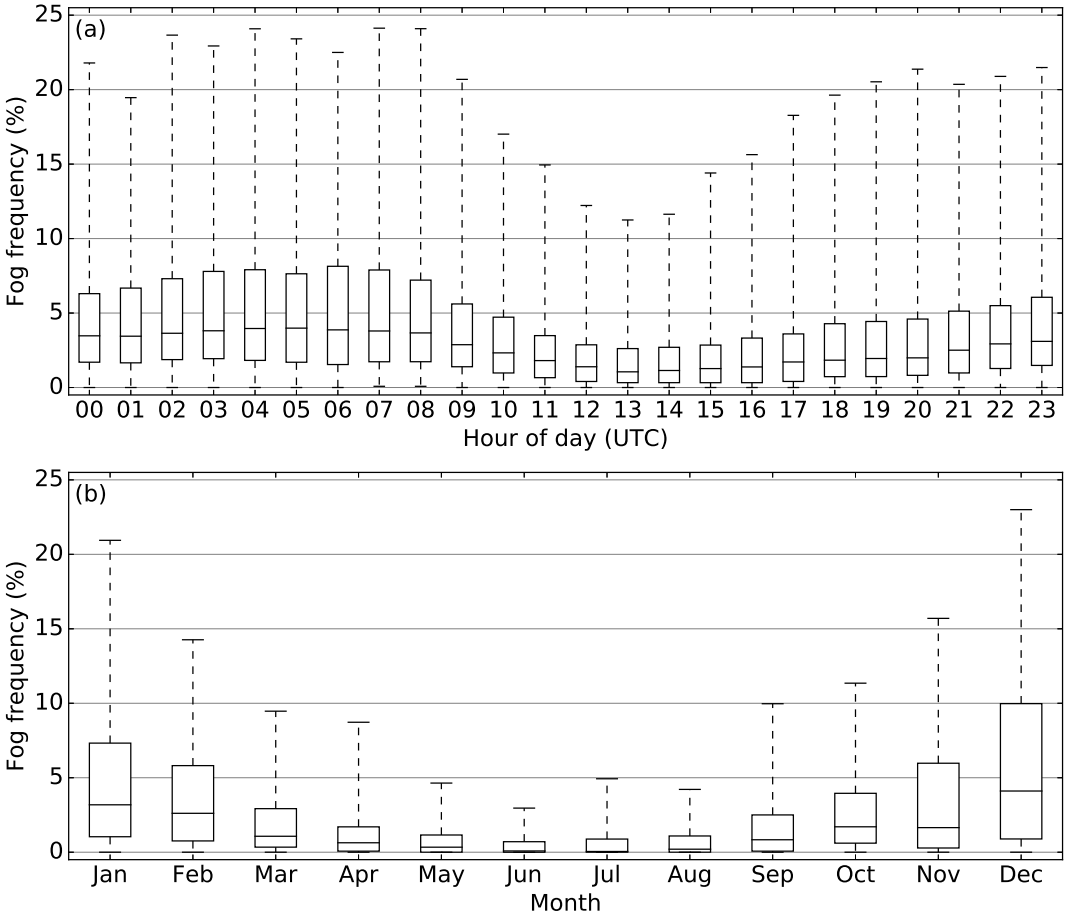

FIGURE 4.1: Diurnal (**a**) and annual (**b**) fog frequency variations measured by all METAR and SYNOP stations used in this study over the complete period from 2006 to 2015. Whiskers mark 5 % and 95 % percentiles, boxes extend from the lower to the upper quartile of the fog frequency data with a line at the median.

0° improves the algorithm applicability significantly. The simultaneous integration of satellite data, terrain information and ground truth station data into the scheme thus makes it possible to capture the strong spatiotemporal dynamics of fog appropriately.

Please note that in this study, fog occurrence is given in raw frequencies (ratio between the number of reports with fog occurrence and the total number of reports at each station). The unit is thus given in percentage without the reference to hours or days. To convert the raw frequencies to fog hours per day, they have to be multiplied by a factor of 2.4.

In Section 4.2 the satellite and ground truth datasets as well as the fog derivation schemes are described in detail. In Section 4.3 the validation procedure and its results are presented. Section 4.4 summarizes the computed fog climatology and section 4.5 gives a final conclusion.

4.2 Data and Methods

In this section, first the general fog retrieval method is described. Then we give an overview of the data sets that were used. After that, we describe the ML model generation and the final fog derivation in detail.

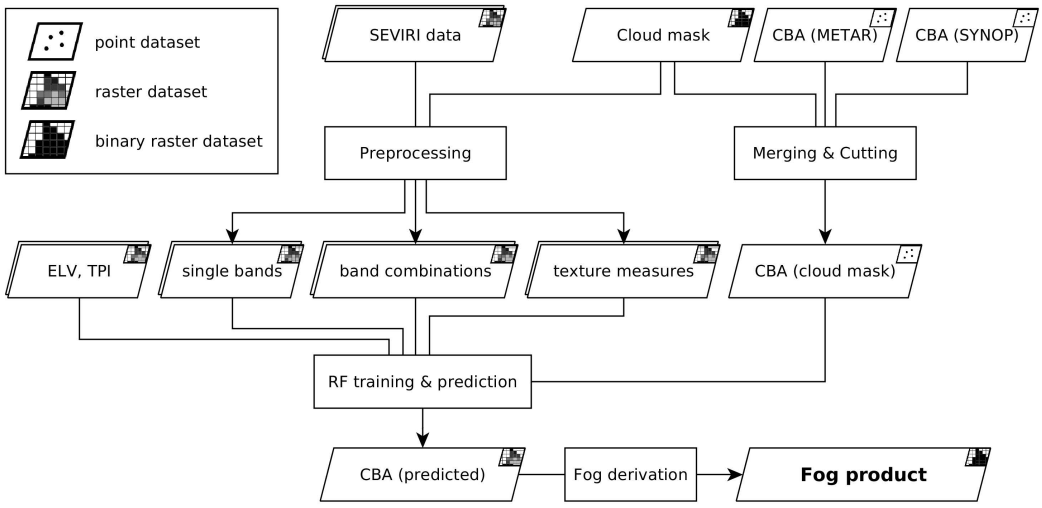

FIGURE 4.2: Schematic view of the fog retrieval scheme. See Section 4.2.1 for a detailed description.

4.2.1 Fog Retrieval Scheme

The aim of this study was to derive area-wide fog information for Europe. To this end, a novel hybrid fog detection algorithm based on a combination of SEVIRI and METAR/SYNOP data was developed. The basic principle of this algorithm is to first derive spatially explicit CBA information in regions with cloud coverage. In a second step, this information is then used to identify potential areas of fog within cloudy regions by finding suitable thresholds for the retrieved CBA values. Here, we have to emphasize again that our definition of fog (VIS $\leq$ 1000 m) includes all conditions that lead to visibilities of less than 1 km at the ground (cf. Glickman (2000)). This includes all types of fog: Advection, radiation, coastal, frontal and orographic fog as well as elevated low stratus being only immersed by higher terrain.

The CBA derivation consists of a hybrid approach (see Figure 4.2). For each MSG scene, a subset of the SEVIRI data, combinations and geostatistical texture features of the original satellite bands as well as static terrain information are used as input features to train a Random Forest (RF) model. The model is trained towards station report data of CBA values that were contemporaneously reported by METAR and SYNOP stations in Europe. It is then used to predict CBA values for all cloud-covered pixels in the domain. An RF model was chosen as ML approach because RFs are computationally cheap and they follow a relatively simple concept which allows for a straightforward interpretation of the generated model output. It was also shown that for comparable classification algorithms the ML model selection had less impact on the quality of the generated output than the selection of the input variables which is why we decided not to use more complex models (cf. Meyer et al. (2015)).

4.2.2 MSG SEVIRI Data

Meteosat Second Generation (MSG) data from the platforms of Meteosat-8, 9 and 10 were used in this study. These geostationary satellites were located over the South Atlantic at or close to 0.0° E/0.0° N during their operational usage. The on-board Spinning Enhanced Visible and Infrared Imager system (SEVIRI) system scans the full hemisphere every 15 min with a sub-satellite resolution of 3 km in 11 spectral bands and an additional high resolution visible (HRV) channel with a sub-satellite resolution

of 1 km. The instrument covers wavelengths from the visible range (0.56 µm) to the thermal infrared (14.40 µm) which allows for a detailed examination of the atmospheric column at these wavelengths (Schmetz et al., 2002). The complete data set between 2006 and 2015 was acquired from the EUMETSAT in level 1.5 format. These data are geolocated and corrected for all radiometric and geometric effects, which makes them suitable for the derivation of meteorological products (EUMETSAT, 2013). However, it has to be emphasized, that the spatial resolution of the satellite defines the size range of detectable fog patches which impedes the detection of small scale fog occurrences, mainly in complex terrain. Also, the recording frequency of 15 min can leave very short fog events undetected.

In addition to the MSG data, the CMA Cloud mask product (V003) provided by CM-SAF which is mainly based on the MSG SEVIRI data was used to filter the domain for cloud-covered pixels (Finkensieper et al., 2016).

4.2.3 METAR Data

Meteorological Aviation Routine Weather Reports (METAR) are standardized weather reports that are reported by airport stations throughout the globe. Their primary use is for aviation safety but they also find application in meteorological tasks. With more than 1300 METAR-reporting stations in Europe, most areas are well covered. Typical report intervals range between 20 min and 60 min which comes close to the temporal resolution of the SEVIRI data. However, measurement instruments and techniques are not internationally standardized which causes variability in the data reliability between METAR stations. Most big airports detect VIS, cloud levels (CL) and CBA using ceilometer and visibility sensors whereas some smaller airfields rely on simple human observation only. As a consequence, from the original data set we filtered land stations with a temporal resolution of less than 30 min and a 24 h report cycle with a complete coverage of the period between 2006 and 2015. After the filtering, 273 stations were left (see Figure 4.3). It has to be emphasized, that due to the time resolution of 20 min to 30 min, it is possible that very short fog events are not recorded by the stations. However, there is currently no better station network available that provides data on CBA and VIS throughout Europe at comparable frequencies. To get equal temporal resolutions in both data sets, METAR reports of the remaining stations were linearly interpolated onto the 15 min intervals of the MSG data. From this data set, we derived the following information:

Cloud base altitude (CBA): When VIS was reported to be below 1 km, CBA was assumed to be 0 m. When VIS was above 1 km and one or more cloud levels were reported, the altitude of the lowermost level was accepted. In METAR reports, all cloud heights are given in feet above ground level (a.g.l.). In order to obtain CBA values in m a.g.l., CBA values were converted to meters.

Fog (binary): As, by definition, fog occurs when visibility decreases below 1 km, the fog flag was set to true when VIS was reported to be below 1 km at the METAR stations. Otherwise it was set to false. It has to be stated, that sometimes low visibilities can be a consequence of precipitation or high aerosol pollution although in most cases they are caused by fog occurrence in Europe.

4.2.4 SYNOP Data

One major disadvantage of METAR stations is that they are bound to airport locations. As airports are mostly situated in flat terrain, this means that areas with

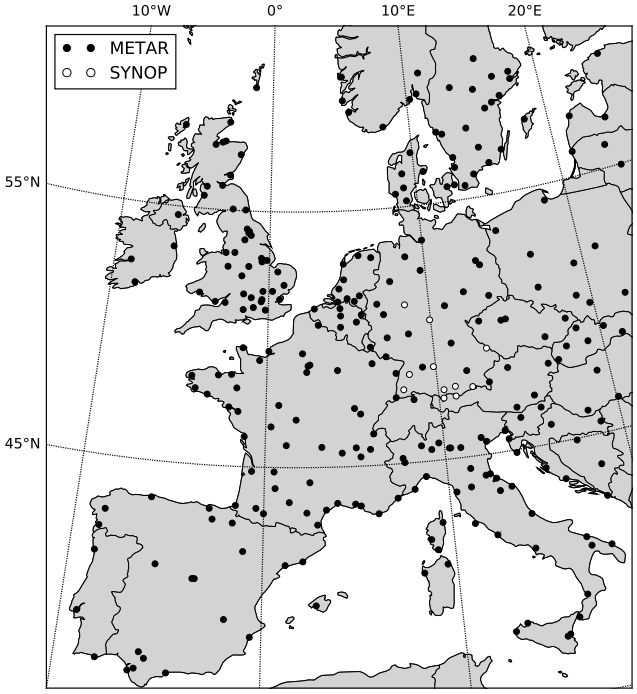

FIGURE 4.3: All 273 METAR (black) and 11 SYNOP (white) stations that were considered in this study after the filtering procedure.

complex topography and higher elevations are underrepresented in the METAR station data. Although the selected METAR stations cover the spatial extent of the study domain sufficiently, they only reach from −5 m above sea level (a.s.l.) to 793 m a.s.l. elevation. To compensate for this disadvantage, additionally eleven SYNOP reporting stations of the DWD were included in this study that lie in higher terrain (see Figure 4.3). Their elevations reach from 705 m a.s.l. to 2964 m a.s.l. and thus take better account of the complete altitudinal gradient of Europe. SYNOP data is reported every full hour and the reports contain VIS and CBA observations. Like the METAR reports, this data was preprocessed by temporally interpolating them onto the SEVIRI scan times over Europe. CBA and binary fog information was also extracted in the same way as for the METAR data.

4.2.5 RF Model Generation

Basic Concept and Implementation

For the prediction of CBA values as part of the fog retrieval, a machine learning model is trained for each scene using pixels with known CBA values (station locations). The model is then used to predict CBA values for all cloud regions within the domain. Here, we use a random forest (RF) regressor, the concept of which was first introduced by Breiman (2001). An RF regressor is an ensemble learning method that is based on a large number of decision trees. Each tree is trained separately by taking a bootstrap sample from the training data set. It is built such that at each node of the tree, a random subset of the original input features is used to split the node. Each node is split by minimizing an error function for the subsets of the respective node. This procedure is recursively repeated until the complete training data set is covered. The RF regression method was chosen because it runs efficiently on large data sets and it can easily handle vast numbers of input features. As the procedure provides internal estimates of the different feature importances it is also well suited to help filtering

for those features that are relevant for the prediction of the target feature. The RF model procedure was implemented in Python using the Scikit-learn package that is provided by Pedregosa et al. (2011).

Preprocessing and Input Features

In previous studies, fog occurrence was mainly linked to atmospheric conditions that favored radiative fog formation which is why FLS is identified in a first step towards fog identification in these studies (see Cermak and Bendix (2011)). However, a preliminary investigation of the ground truth and satellite data sets used in this study showed that only 33.2 % of all fog occurrences reported by the ground truth stations could actually be identified as covered by FLS from the satellite's perspective. There are three reasons for this: (1) Multiple cloud layers obscure the satellite's view: 60.1% of all pixels for which FLS was reported by the respective ground truth station were covered by higher opaque clouds which prevents proper FLS detection from space. (2) Fog patches smaller than 9 km^2 are too small for the MSG resolution and cannot be captured by the satellite accordingly: 23.2% of all pixels with fog occurrence reported by the respective ground truth stations could not be identified as cloudy by the satellite. (3) Orographic fogs and other fog types that are not dependent on local radiative cooling are not dependent on FLS occurrence and can form below other cloud types. Due to these reasons, a meaningful fog climatology can only be generated by considering all cloud-covered pixels. Contrary to previous approaches, a preliminary filtering of FLS pixels was thus not conducted in this study.

A tuning data set was extracted from the original data so that the RF models could be tweaked in order to deliver the best possible results: In a first step, reflectances as well as brightness temperatures of each single band were extracted from the raw data. After that the SEVIRI and METAR/SYNOP data sets were merged by assigning each CBA report from each METAR/SYNOP station to the corresponding pixel values of each band of the respective SEVIRI scene. This was done for all 342'328 available SEVIRI scenes between 2006 and 2015. As we were only interested in regions that are covered by clouds, the merged data set was reduced to samples where the CM-SAF cloud mask product (Finkensieper et al., 2016) showed cloud-covered pixels. From this global data set a sample data set was extracted by selecting 100 random scenes from each single month which resulted in a data set of 11'993 valid MSG scenes with their corresponding METAR/SYNOP reports which again yielded in 395'560 samples in the final tuning data set.

As the combination of different bands proved to be helpful to detect FLS e.g., Bendix and Bachmann (1991), it seemed reasonable to also provide the models with band combinations. Although this leads to partial redundancy in the input features, it helps to highlight otherwise unapparent patterns in the data. Cloud top height (CTH), liquid water path (LWP) and cloud phase (CP) can be used as proxies towards CBA estimation. As these microphysical properties can be derived from band combinations between the water vapour bands and the middle and thermal infrared bands, the five most important combinations between these bands were included in the model data (cf. Kühnlein et al. (2014), Thies et al. (2008)).

Still, spatial information, such as the heterogeneity within cloud patches, is not covered by only looking at single pixel values. Information about spatial variance within a given raster data set, however, can improve the output of ML approaches. This is exploited in many ecological studies, e.g., Gloaguen et al. (2008) or Schulz et al. (2017). As spatial information in the form of standard deviations within cloud patches is also included in the FLS detection algorithms of Cermak and Bendix (2008)

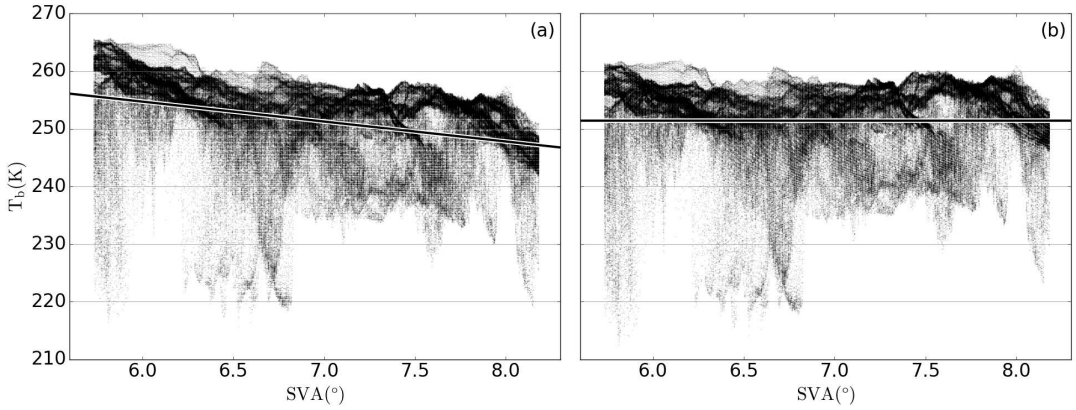

FIGURE 4.4: SVA correction at the example of the 7.3 µm water vapor band on September 5, 2012 05:00 UTC. (a) shows the original band values, (b) shows the corrected band values plotted against SVA. The black line marks the linear regression for each case.

and Egli et al. (2017), it seemed promising to also provide this information to an ML model to improve the prediction of CBA values. Thus, several geostatistical texture features of the 5 × 5 pixel surroundings of each METAR station were provided for the model training. These measures include standard deviations (STD), variograms (VAR), madograms (MAD) and rodograms (ROD) of each band and cross variograms (CV) as well as pseudo cross variograms (PCV) of each band combination. Please refer to Equations (4.1)–(4.5) in the Appendix for a detailed mathematical description of the listed features.

Although MSG SEVIRI level 1.5 data includes basic corrections, two effects caused by the satellite's geostationary orbit have to be considered: Variations in the solar position and in the satellite viewing angle (SVA).

Seasonal and diurnal variations in the solar cycle influence the responses of the MSG bands, especially those in the visible and near infrared range between 0.6 µm and 3.9 µm. However, these variations have to be considered in the model training only if the model should be able to predict CBA values for multiple time steps with changing solar positions. In our approach, however, an independent RF model is trained separately for each scene. As intra-scene solar angle variations are relatively small due to the small domain size, the sun's position was not considered as an input feature to the model training.

SVA reflects the distance between the earth's surface and the satellite sensor. This also means that the atmospheric column and therefore the amount of absorbing gases increase with growing SVA. This causes distortions in the value distributions of the channels within the domain. Unfortunately, SVA cannot be used as an input feature for the RF training in order to correct this effect, because its high spatial autocorrelation would lead to artifacts in the model prediction results around the station pixels. Instead, the scenes were cleared from the influence of the SVA before handing them over to the RF training. An empirical approach was chosen: First, average reflectance and brightness temperature values were calculated from all 11'993 scenes of the tuning data set separately for each band. Then, regressions were calculated between the averaged bands and the corresponding SVA values. This was done separately for each band. The regression coefficients determined in this way were then applied to each scene using Equation (4.7) to remove the SVA influence from the data before handing them over to the RF training. An example of the SVA correction is depicted in Figure 4.4.

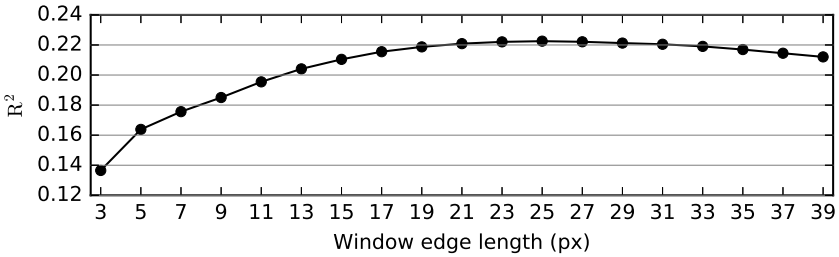

FIGURE 4.5: Correlation coefficients between TPI and CBA for different window sizes.

MSG Bands	Derived Data	Static Data
VIS 0.6	ΔT 7.3–12.0	ELV
VIS 0.8	ΔT 8.7–10.8	TPI
IR 1.6	ΔT 10.8–12.0	
IR 3.9	ΔT 3.9–7.3	
WV 6.2	ΔT 3.9–10.8	
WV 7.3	STD (all bands)	
IR 8.7	VAR (all bands)	
IR 9.7	MAD (all bands)	
IR 10.8	ROD (all bands)	
IR 12.0	CV (all band comb.)	
IR 13.4	PCV (all band comb.)	

TABLE 4.1: All input features used for RF tuning. MSG bands are denoted with their central wave length in µm.

As CBA is substantially defined by the condensation level which is closely linked to the planetary boundary layer along warm fronts which again is closely linked to the underlying topography, terrain elevation (ELV) as well as the Topographic Position Index (TPI) were also added to the training data. TPI values are highly dependent on the window size used for their calculation. Thus, the correlation coefficients of a linear correlation between CBA and TPI were calculated for different window sizes using the tuning data set (see Figure 4.5). A window size of 25 pixels edge length was used for the final RF models since window sizes with an edge length of 25 pixels resulted in a maximum R^2 value of 0.22.

ELV was extracted from the WorldClim DEM generated by Hijmans et al. (2005). Please find a detailed description of the calculation of the TPI in Equation (4.6) in the Appendix. All input features that were used during the RF model generation are listed in Table 4.1.

Feature Selection

In order to reduce the "curse of dimensionality" (Bellman, 1961) and to obtain the largest possible information content from the original data while reducing noise artifacts due to irrelevant information, the features that are most relevant for CBA derivation were extracted from the original data set. To this end, a recursive feature elimination (RFE) was conducted for each scene of the tuning data set starting from randomly varying 20 feature subsets. In each step of this iterative procedure the original number of input features was reduced by one element and the model quality was calculated using an out-of-bag (OOB) R^2 score. The OOB R^2 score is calculated on the basis of the left-out data during model fitting and is similar to the outcome of a cross-validation (Breiman, 1996). This iterative procedure was repeated for each scene until only one feature remained. The results are depicted in Figure 4.6.

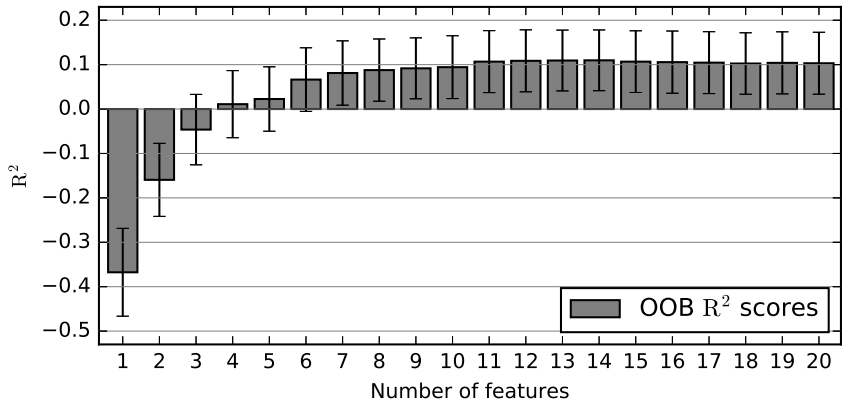

FIGURE 4.6: Results of the feature elimination. Bars denote averages, error bars denote standard deviations of the R^2 scores reached during all runs for each scene in the tuning data set.

Average RF model performances were almost constant with R^2 scores around 0.10 for feature counts from 74 down to 14. On average, the best model performance with an R^2 of 0.11 was reached at 14 remaining features in the tuning data set. Removing further features from the set resulted in worse scores. Thus, for the final RF model fitting procedures, the 14 features with the highest average rankings were filtered from the original set of 74 features and used as training input for the final RF models (see Figure 4.7).

The plot shows that the terrain features were identified as having the highest explanatory power in terms of CBA derivation. This shows that the models identified the close link between condensation level—which defines the lower cloud boundary—and terrain elevation. The high ranking of TPI can probably be attributed to the fact that TPI can help identifying situations where the condensation level does not closely follow the terrain elevation but is shifted towards or away from the surface due to local ridges and troughs. However, the standard deviations of these two features are relatively high compared to the other features. This is probably due to the fact that TPI and ELV only provide explanations for the cloud height in LS situations where only the higher elevations reach into the cloud layer. This distinction is not necessary for plane-filling fog fields and the two terrain features cannot contribute to the CBA retrieval in these cases.

The bands in the thermal infrared, especially at 3.9 µm and 10.8 µm make up the biggest part of the remaining features. This is not surprising as the thermal radiance emitted by the surface and by clouds which is recorded by these bands can be used as a proxy for cloud top altitudes. This is also the reason why these bands are vastly used in many cloud altitude derivation algorithms (Hamann et al., 2014). Additionally, due to the different emissive behaviors of small cloud droplets in the 3.9 µm and 10.8 µm bands, the differentiation between high and low lying stratus layers is possible which can be used as a first proxy towards the final CBA estimation. As the ΔT 3.9 – 10.8 combination is placed on rank 4 it seems that the models were indeed able to use this information for the CBA prediction.

The water vapour bands at 6.2 µm and 7.3 µm do also make a major contribution to the CBA estimation. This is surprising in that both bands are insensitive to the lower atmosphere due to the strong water vapor absorption in the upper layers. The bands are thus not able to capture ground-level atmospheric conditions which are expected to be most relevant for the derivation of low cloud base altitudes. As the bands

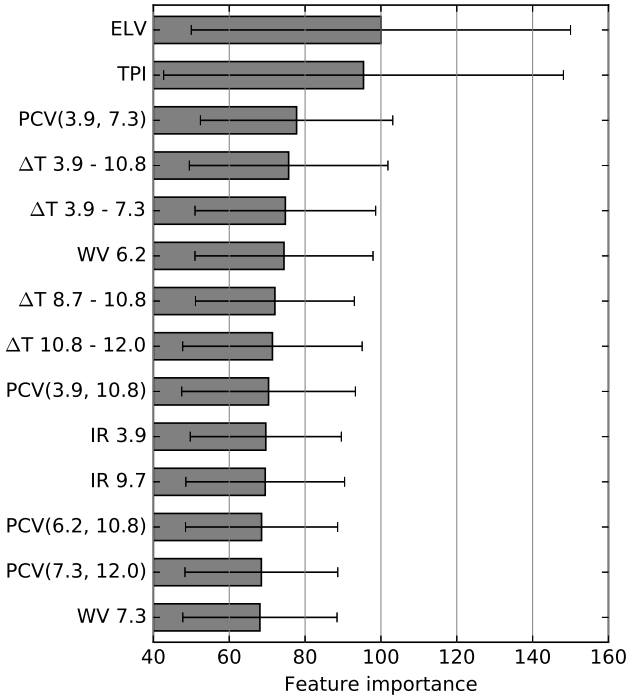

FIGURE 4.7: Average feature importances and standard deviations relative to ELV (100) of the 14 input features used for the final RF model generations. Feature importances were calculated for all 11'993 sample MSG scenes from the tuning data set using all input features.

are, however, able to detect upper atmospheric circulation patterns, they can bring information about the large-scale synoptic situation into the model which could indeed help to deduce low cloud altitudes. Due to their ability to derive humidity information from the middle and high troposphere, both bands can also be used in combination with the IR window bands to differentiate between convective and stratiform clouds (Tjemkes et al., 1997). This information is crucial for the final fog retrieval as low visibility due to fog is more likely to happen below a stratiform cloud layer than below convective systems (e.g., as a result of ground touching LS in anticyclonic conditions or as frontal fog along warm fronts in winter) .

On the other hand, the visible bands at 0.6 µm and at 0.8 µm ranked lowest on the feature list and were not included in the final model training. This was expected as they only capture the reflectance of sunlight and no thermal radiance emittance. Though this helps to delineate cloud entities during daytime, it does not provide information about their altitude via brightness temperature estimations.

In general, the models chose to use band combinations and texture measures over the use of single bands. It is also obvious that the models preferred to use two bands in combination as the only texture measure that was used is the pseudo cross variogram. All this suggests that the prior processing of the raw data was indeed helpful in uncovering otherwise obfuscated patterns.

Parameter Tuning

Following the approach of Kuhn and Johnson (2013), the total number of decision trees (n_t) and the number of input features used at each node (n_n) were tuned to get optimal settings for the RF model. This was done in an iterative procedure with n_t ranging from 50 to 300 and n_n ranging from 1 to 5. For each parameter set RF

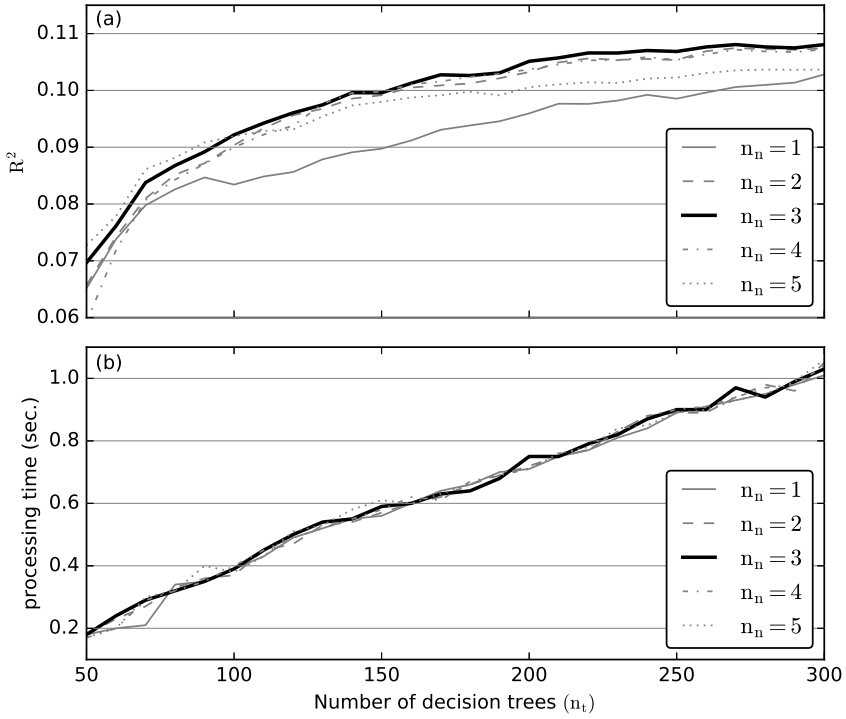

FIGURE 4.8: RF parameter tuning results for n_n between 1 and 5 and n_t between 50 and 300. (a) shows R^2 scores and (b) shows processing times per MSG scene for the different RF parameter settings.

models were fitted to the tuning data set and the OOB R^2 score was calculated. The parameter tuning procedure was run on an Intel Core i7-5600U CPU @ 2.60GHz × 4. The results are shown in Figure 4.8.

The model performance strongly increased in the beginning and then slowly approached an upper limit of about 0.11 after $n_t = 250$. With $n_n = 1$ the score values were significantly lower than when more features were considered at the nodes. On the other hand, n_n values larger than 3 didn't increase the prediction performance. With n_t values growing above 250 the performance increase began to stagnate whereas processing times linearly grew further (see Figure 4.8 (b)). As a compromise between prediction accuracy and processing time, a cut has been made at $n_t = 250$ and n_n was set to 3 in the final RF models that were used for the fog retrieval scheme.

Application of the RF Models and Fog Derivation

Feature elimination and parameter tuning procedures were carried out using the tuning data set to determine both the final input feature set and the best model parameters. These settings were than applied to the processing scheme of each single MSG scene. Ground truth station observations of CBA that were temporally interpolated onto the MSG scan times were merged with the satellite information and used as model input (see Figure 4.9 (a) for an example). Within the scope of this study, only the SYNOP stations in Germany were available as representatives for higher terrain and exposed locations. As these locations showed significantly lower CBA values than the remaining stations and due to the spatial autocorrelation in the input features, a simple application of the RF models including all input features would have led to an adequate prediction of CBA values within Germany but a strong overestimation of CBA values in the rest of the domain. Therefore, the RF application had to be

divided into two steps for each scene. First, a model was trained using all available METAR stations within the respective scene that were covered by clouds. As input feature set only the satellite bands and their derivations (geostatistical texture features as well as band differences) were used. The CBA values predicted with this model are therefore based only on the information content of the satellite bands. A second model was then trained using the output of the first model and TPI as well as ELV as input features. Using this two-step approach, the influence of the terrain could be incorporated into the model predictions without spatially over-fitting CBA values to Germany.

Because of their structure, RF models are able to predict almost perfectly the target variables of samples that appear in both, the training data set as well as in the prediction data set. The resulting overfit of the model to the station locations can be prevented by a slight change in one of the independent variables of the sample which is why for each model training procedure, instead of taking the actual pixel values of the station locations, the average of their direct pixel neighbours were used. Although the data accuracy is artificially impaired by this step, it is necessary to prevent the algorithm from overfitting to the station pixels. Figure 4.9 (b) shows an example output of the RF model predictions for 15 November 2011 00:00 UTC. The model was clearly able to distinguish high clouds over southwest France, Spain and the Balkans from fog and LST patches in England, Germany, northern Italy and the Benelux countries.

As, by definition, fog occurs where a cloud touches the ground, pixels with values equal to or below zero could theoretically be classified as pixels with fog occurrence. However, reported CBA values are never negative and thus the RF model predictions are never smaller than zero although the theoretical condensation level could actually be considerably lower than the terrain elevation. As only the pixels of fog reporting stations can have CBA values of zero, this would lead to a strong underestimation of fog occurrence for non-station pixels without an application of a threshold to the CBA predictions. The threshold values were derived by gradually lowering the layer of predicted CBA values and calculating Heidke Skill Scores (HSS, see Equation 4.12) of derived fog occurrences (CBA ≤ 0 m) and observed fog occurrences (VIS ≤ 1000 m) for each step. The shift resulting in the highest HSS was accepted as threshold for the final fog derivation. The thresholds fluctuated around a median of 70 m and remained below 247 m in 90 percent of the cases. The resulting output of this step is shown in Figure 4.9 (c) together with fog observations of the ground truth stations. By applying the threshold value to the CBA predictions, the fog retrieval scheme provides a reasonable fog distribution for the study domain. The Po valley and large parts of southern Germany and northeastern France were marked as fog-covered. In particular, small-scale areas such as the Upper Rhine Valley were recognized as fog-free (elevated LST), while the slopes of the nearby Black Forest are suspected to be covered in fog. However, large-scale structures that can be anticipated from the station data are also visible in the product: While the southern edge of the FLS patch, which stretches from the Netherlands through northern Germany to Poland, was identified as ground-touching, the northern areas within this patch are marked as fog-free. This is consistent with the station reports which indicate that the condensation level only reaches the surface at the southern end of the patch.

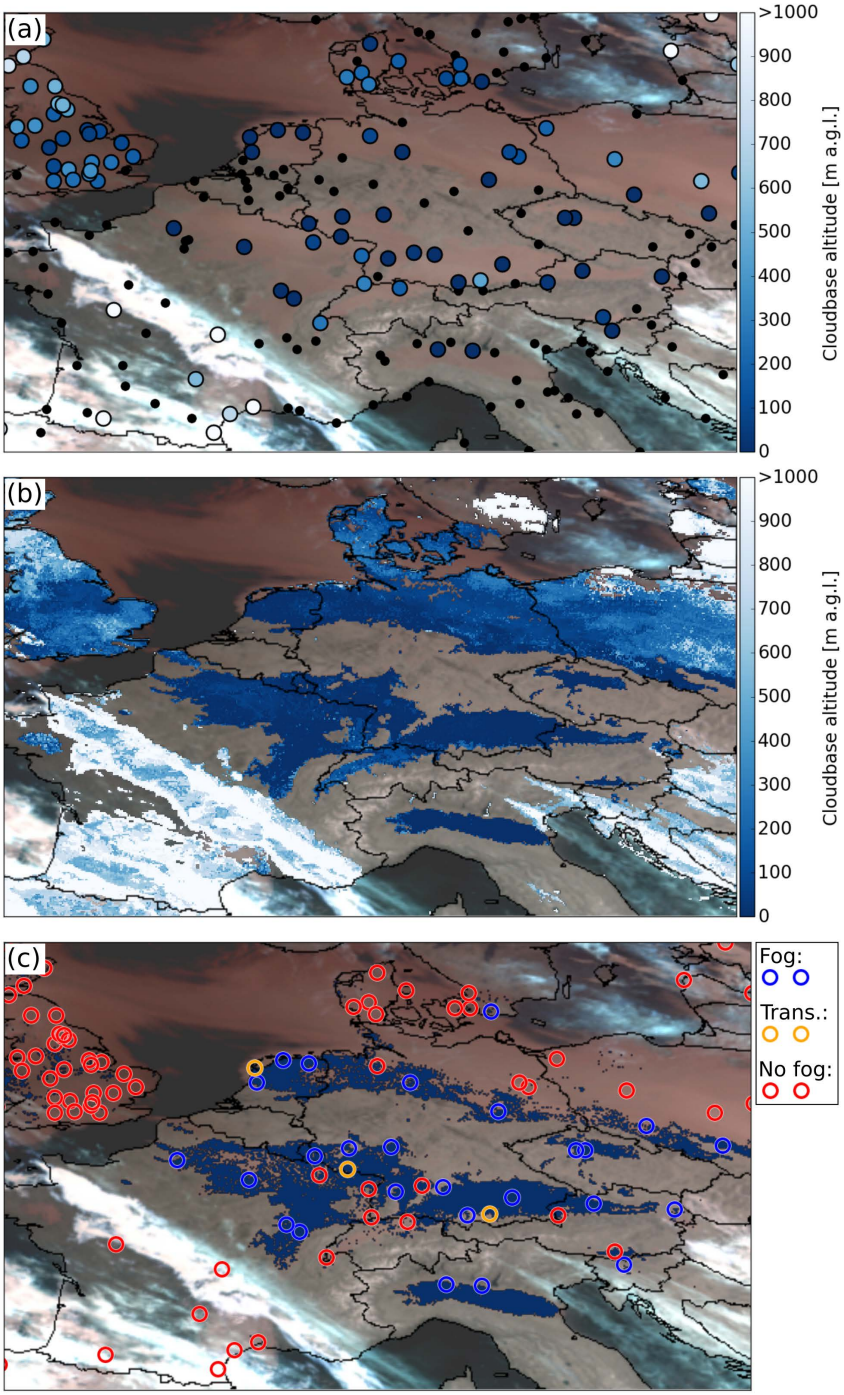

FIGURE 4.9: The basic steps of the fog retrieval scheme at the example of November 15, 2011 00:00 UTC. The background of each subplot is an IR composite of the 3.9 µm, 10.8 µm and 12.0 µm band. The domain is cropped to Central Europe in order to make small scale patterns visible. (a) shows a combination of ground truth and satellite data. Circles mark positions of CBA reporting stations. Colors denote observed CBAs. Black dots indicate that no CBA was measured at the given time. (b) displays RF predictions for all cloud-covered regions above land. (c) presents the final fog derivation (blue areas) and station observations (colored circles). Blue circles mark stations with fog occurrence, yellow circles mark stations where fog formed or dissipated during the recording of the satellite scene and red circles mark stations where no fog was observed.

4.3 Validation

To get an overview of the reliability of the fog retrieval scheme and the validity of the produced data set, an extensive validation procedure was conducted. Here we chose to validate the derived CBA data and the final fog product using the METAR/SYNOP stations that were also taken for the RF training in a cross validation approach.

At this point it has to be stated, that satellite products are not directly comparable to ground truth data for a number of reasons: Instabilities in the satellite's orbit cause parallax shifts which lead to ambiguities in the assignment of ground-truth data locations to pixel coordinates. Also, different temporal resolutions in both datasets result in timing uncertainties and multiple cloud layers as well as subpixel effects resulting from the satellite resolution at margins of fog patches additionally impede the comparison between satellite and ground truth data. All these insufficiencies cause ambiguities in the comparison process. However, there are no other reliable data sources for fog occurrence that provide better information on fog distribution on a continental scale.

The validation was conducted using a leave-location-out (LLO) cross validation procedure. By using an LLO approach, it is guaranteed that the validation results are not corrupted by a possible overfit of the model outputs to the station pixels and that the validation results are generally applicable to the complete study domain. The LLO procedure was carried out in such a way that the RF models were trained for each scene of the tuning data set using the input data of all but one station. The CBA prediction results were then compared to the observations of the left-out station. Fog was derived from CBA values by application of the threshold technique explained in Section 4.2.5. As the LLO cross validation procedure is very time-consuming and computationally expensive, only 10 randomly drawn stations per scene were used for the fog derivation. Observed and predicted CBA and fog values were then compared between these stations. This procedure was repeated for all scenes in the tuning data set. As the models were only trained for cloud covered regions in the CM-SAF cloud mask, only these pixels were included in the validation procedure. This resulted in a total of 118'145 samples, which were included in the validation procedure. Based on the generated data set, the following validation measures were calculated to get an overview of the quality of the product.

Average absolute differences between predicted and observed CBA values were calculated for each sample and MAE (Mean Absolute Error) values were computed for each station. PC (Percentage Correct), POD (Probability Of Detection), FAR (False Alarm Ratio) and HSS (Heidke Skill Score) were calculated for the fog predictions accordingly. Please refer to Equations 4.8–4.12 in the Appendix for a detailed description of these validation measures. While PC, POD and FAR address very specific features of the fog data set only, the HSS compares the forecast accuracy of the product to the forecast accuracy of a random guess procedure. This gives an overall indication of the validity of the algorithm output, independent of the fog frequency. The HSS ranges between -1 and 1. Negative values indicate that the model output is worse than a random guess, a value of zero means no skill increase compared to a random guess while a perfect prediction obtains a HSS of 1. All verification measures were calculated for all cloud covered samples in the tuning data set. Diurnal and annual variations in the verification measures are plotted in Figures 4.10 and 4.11.

MAE values fluctuate around an average of 298 m with minima of 251 m at 08:00 UTC and 267 m in February and maxima of 340 m at 23:00 UTC and 325 m in July (see Figures 4.10 (a) and 4.11 (a)). The probable explanation for the better performance in winter is that LST layers are more dominant in winter than in summer. Therefore,

the heterogeneity within the clouds is smaller in winter which in turn simplifies the derivation of cloud base altitudes.

As already mentioned in the introduction, the validation data set also shows that fog frequency is considerably higher at night and in winter due to the general decrease of the condensation level than during the day and in the summer months. Daily and annual frequency variations are stronger in the predicted values than in the observation data. Higher amplitudes in winter and at night in the predictions most likely are a consequence of fog occurrences in the lowlands are more dominant in these times. It is more difficult for the retrieval scheme to correctly derive the spatial extent of these fogs, since they occur in flat terrain and vertical shifts of the thresholds used for fog derivation have a stronger effect here. This leads to a slight overestimation in the generated product during these times. The daily maximum is reached around 04:00 UTC with 8.3 % in the observations and 9.5 % in the predictions. Fog frequency then decreases steadily as the day progresses and reaches its minimum of 2.4 % in the observations and 2.1 % in the predictions at 15:00 UTC. Afterwards it starts to rise again steadily. The annual curve behaves similarly with a maximum of 8.5 % in December (observations) and 11.1 % in January (predictions). Minima occur in June with 2.8 % (observations) and 2.1 % (predictions). With an average value of 95.6 %, PC is always on a high level. However, with maxima of 98.1 % at 15:00 UTC and 97.5 % in June as well as minima of 93.2 % at 03:00 UTC and 92.7 % in January, the PC curve is almost an inversion of the fog frequency curve in both, diurnal and annual averages. This is due to the fact that significantly fewer fog events have to be "hit" during the day and in the summer months which drives up the PC due to the large number of true negatives (see Figures 4.10 (b) and 4.11 (b)).

The shape of the POD curve shows that the probability of correctly predicting a fog event is higher in winter with a maximum of 0.74 in December, despite the high PC values in summer. Contrary, in summer POD generally decreases down to a minimum of 0.44 in June. The poor performance of the POD in summer is mainly due to the fact that the absolute number of fog events is very low in the warm months which makes a “hit” more difficult. The FAR values behave similarly to the POD values at a lower average level of around 0.41. They also show a less pronounced daily and annual amplitude with maxima of 0.48 at 11:00 UTC and 0.49 in June as well as minima of 0.34 at 20:00 UTC and 0.35 in August. This means that the relative rate of false positive predictions stays on a moderate level for each period. Both, diurnal and annual HSS curves show very little variation. With an average of 0.58 the scores consistently remain on a high level and thus indicate the good performance of the scheme throughout the complete time period (see Figures 4.10 (c) and 4.11 (c)).

In addition to the direct validation procedure on a 15 min level we compared average fog frequencies of the product to average station observations of the complete time period from 2006 to 2015 (see Figure 4.12). We can see that the product matches the observations with a bias of −0.2 % and an absolute average error of 0.7 % very well at the METAR stations. At the SYNOP station pixels, however, the bias is much stronger with −10.4 % which leads to an underestimation of fog frequencies in mountainous regions. This is mainly due to the fact that the scheme can only flag a pixel as covered by fog if the satellite was able to identify it as cloudy beforehand. All fog patches with sizes in the subpixel range thus cannot be detected by the approach and are therefore not included in the statistics. This effect is stronger at SYNOP stations because the probability of subpixel fog patches is higher in the heterogeneous terrain of the SYNOP stations due to the large differences in terrain height within small distances. Nevertheless, with an average absolute error of 1.1 % over all stations, we can assume that Figure 4.13 gives a reasonable overview of the general distribution

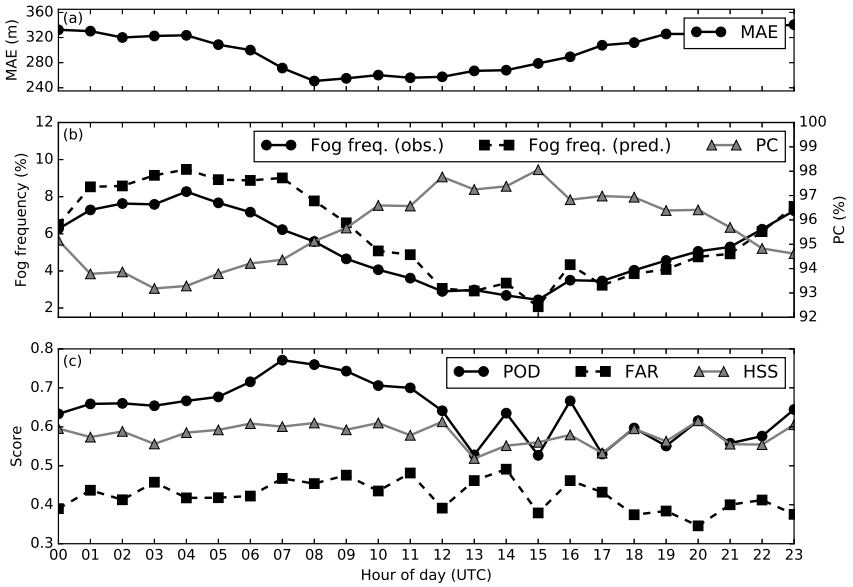

FIGURE 4.10: Hourly variations in LLO cross validation measures averaged over all years and all ground truth stations: (a) MAE of CBA predictions, (b) observed and predicted fog frequencies as well as PC, (c) POD, FAR and HSS.

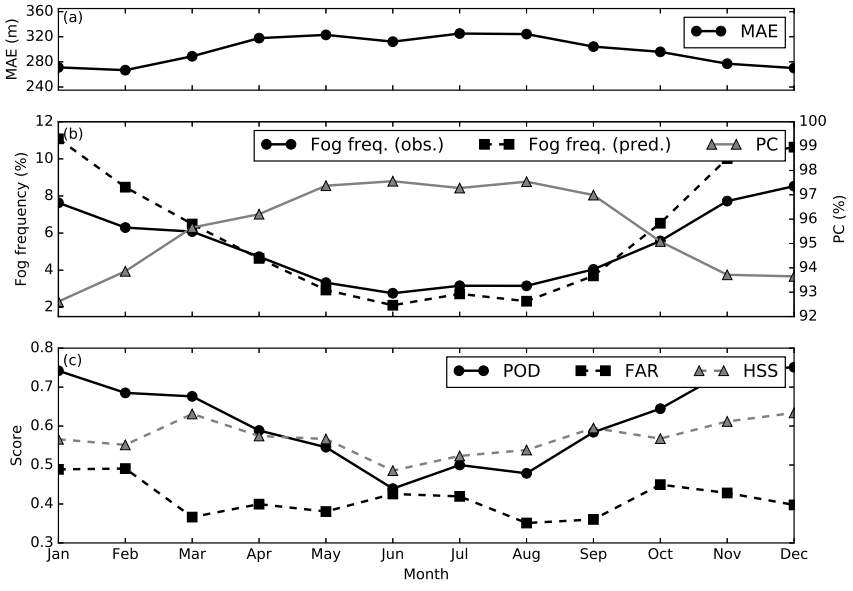

FIGURE 4.11: Monthly variations in LLO cross validation measures averaged over all years and all ground truth stations: (a) MAE of CBA predictions, (b) observed and predicted fog frequencies as well as PC, (c) POD, FAR and HSS.

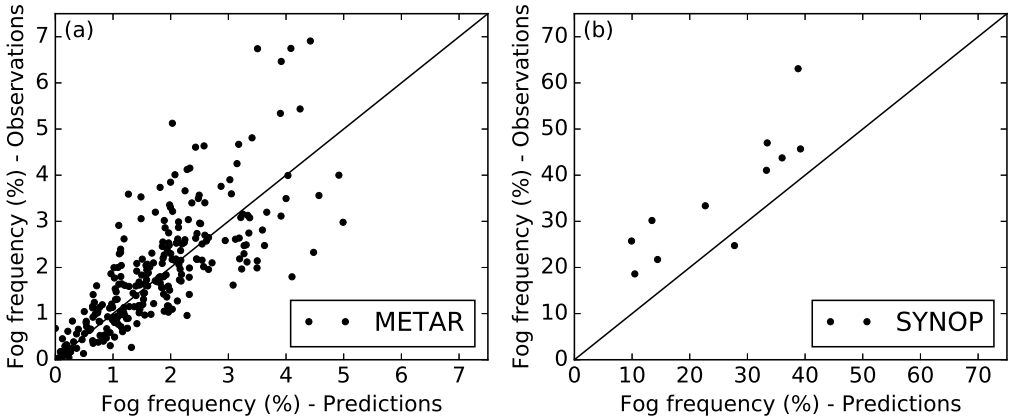

FIGURE 4.12: Predicted vs. observed fog frequencies for all (a) METAR and (b) SYNOP stations. Values are averaged over the complete period from 2006 to 2015.

Measure	SOFOS	Hybrid
PC	0.99	0.95
POD	0.52	0.61
FAR	0.66	0.41
HSS	0.41	0.58

TABLE 4.2: Verification measures of SOFOS and the hybrid scheme applied in this study.

of the fog frequency in Europe.

To compare the performance of our approach with the performance of existing physics-based methods, the same verification measures were calculated from the confusion matrix of SOFOS (Satellite-based Operational Fog Observation Scheme) presented in Cermak and Bendix (2011). The results are listed in Table 4.2 together with the verification measures of this study. While SOFOS was able to reach a higher PC value (0.99 vs. 0.95), the hybrid scheme performed better in the other verification measures. The high PC score in SOFOS results from a large number of correctly predicted non-fog situations which is associated with generally low fog frequencies in the investigated time period. With POD values of 0.61, however, the hybrid scheme was able to correctly detect fog occurrences with a higher precision than SOFOS (0.52) while fog-free situations were less often misclassified as foggy with FAR values of 0.41 for the hybrid scheme and 0.66 for SOFOS. Overall, this results in a better performance of the hybrid approach with an HSS of 0.58 compared to 0.41 for SOFOS. It has to be emphasized, however, that the results of both schemes cannot be directly compared to each other as both were evaluated for different time periods. Also, SOFOS is designed for radiation fog conditions at sun elevations above 10° whereas the hybrid approach is generally applicable for all sun elevations and fog types.

In summary, it can be said that the scheme was able to detect fog with higher precision than previous physics-based approaches despite the problems that the satellite-based hybrid approach has to deal with and although RF model scores of 0.11 were relatively low. With slight variations in the daily and yearly cycle, the validation measures consistently indicate satisfactory data quality in the final fog product.

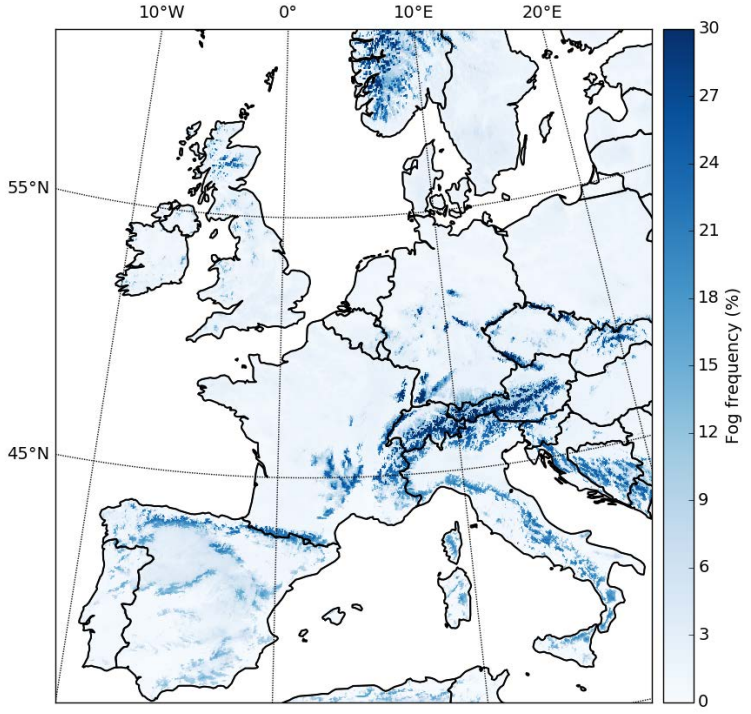

FIGURE 4.13: Fog frequency averaged over the complete period from 2006 to 2015.

4.4 Fog Climatology

The fog retrieval scheme was applied to each of the 342'328 available SEVIRI scenes from the period between 2006 and 2015. The resulting fog product was averaged over all scenes (see Figure 4.13). It is clearly visible that the fog frequency distribution is highly dependent on the terrain and that there can be strong changes in the fog frequency within a few kilometers, especially in and around mountain ranges. Lowlands and coastal regions show low fog frequencies between 0 % and 5 %, while mountainous regions with a maximum of 39 % in the Central Alps are dominated by much higher fog frequencies.

At this point it has to be emphasized that any kind of ground touching cloud is interpreted as a fog event by the algorithm. The map thus does not only indicate radiation fog frequencies that are associated to FLS patches but all kinds of fog types. Also, since fog only prevails where an FLS patch actually touches the ground and since CBAs are relatively homogeneous within an FLS patch, fog usually only occurs at the margins of FLS patches where the terrain height increases. The fog map is therefore almost an inverted version of the FLS map that was presented in Egli et al. (2017) where lowlands and river valleys are prone to higher FLS frequencies than their neighbouring mountain ranges.

Another peculiarity is that the fog map reveals the main flow direction of low and humid/saturated air masses in mountain ranges. Orographic fog forms as a result of condensation due to uplift of humid air at windward mountain slopes. Cloud fog results from stratiform clouds that are formed and transported along warm fronts towards the mountain range. Both fog types form at the windward side of the mountain ranges whereas the leeward side is fog free. This causes higher fog frequencies at the side of the main wind direction. In the Cantabrian Mountains, the Pyrenees, the Alps and the Scandinavian mountains this results in higher fog frequencies on the north

and west slopes of the mountain ranges.

In order to investigate the spatial differences in the diurnal and annual variations in fog frequency, anomalies were computed for individual hours and seasons (see Figures 4.14 and 4.15). The anomaly values were calculated as the differences between each hourly and each seasonal 10 year average and the global mean from Figure 4.13. Values are given in percentage points (pp)—the arithmetic difference of two percentages.

Figure 4.14 (a)–(d) give a good overview of the daily fluctuations in fog frequency. During the first half of the night there is a slightly higher fog frequency than on average in lower altitudes but a strong negative anomaly above 900 m. The distribution pattern changes during the night so that in the morning hours around 06:00 UTC high fog frequencies are visible at all elevation levels. This is due to the fact that the conditions for radiation fog, which forms in lowlands and valleys during anti-cyclonic weather situations, generally improve over night. Additionally, the condensation level decreases as a result of the general drop in temperature until sunrise, which means that the mountain slopes and higher mountainous regions are also covered in fog more frequently as a result of cloud base lowering fog.

After sunrise the situation changes drastically. Lowlands below 1000 m show a negative anomaly at 12:00 UTC whereas the higher regions, especially the Alps and the Pyrenees are covered in fog far more frequently now. However, low visibility situations are not a result of radiative cooling in these cases, but more likely due to orographic fog formation. This happens when humid air is lifted upwards at the windward side of the mountain ranges during cyclonic weather conditions which leads to condensation due to adiabatic cooling. At 18:00 UTC the fog frequency is at a low in both the higher regions and the lowlands. Frequencies increase only slowly as the night progresses, starting in the lowland.

As expected, the maps of the seasonal anomalies (Figure 4.15 (a)–(d)) show that fog frequencies increase in autumn and especially in winter. On the other hand, in spring and even more so in summer, there is a substantial decrease in the general fog frequency. Terrain elevations around 1500 m show strongest variations whereas the lowlands below 500 m and mountainous regions above 2500 m are characterized by much less seasonal variation (Figure 4.15 (e)–(h)).

In winter, the strongest positive anomalies occur in the uplands and at the slopes of high mountain ranges. This applies to the Cantabrian mountains, the Pyrenees and the Sierra Nevada in Spain, the Cévennes, the Vosges and the Jura in France, the Scandinavian mountains, the Highlands in Scotland, the Black Forest, the Swabian Alb, the Harz Mountains, the Rhön, the Ore Mountains and the Bavarian Forest in Germany, the High Tatras, the Carpathians, the Dinaric mountains in Eastern Europe, the Apennines in Italy and the Alps. The reason for this is the general presence of more FLS and the lowering of the condensation level during winter. As a result, mountain slopes and areas of relative elevation often protrude into the extensive stratus fields in the lowlands. Fog frequencies in the lowlands are also increasing in winter. Since the condensation level does not reach the ground level of the lowlands in most of the cases, the difference to the long-term mean is, however, smaller here. Then again, some lowland regions show vastly more fog occurrence in winter. The northern and southern flanks of the Po valley are such an exception. With positive anomalies of up to 9.1 pp fog is much more prominent in these regions. The Po valley has long been known for the frequent occurrence of fog and has therefore already been the subject of several fog research projects (see Bendix (1994), Fuzzi et al. (1992), Mariani (2009), Wendisch et al. (1998)). Due to the orography of the valley, which is enclosed by the Alps and the Apennines on three sides, strong temperature inversions can develop here during anticyclonic conditions. These temperature inversions promote

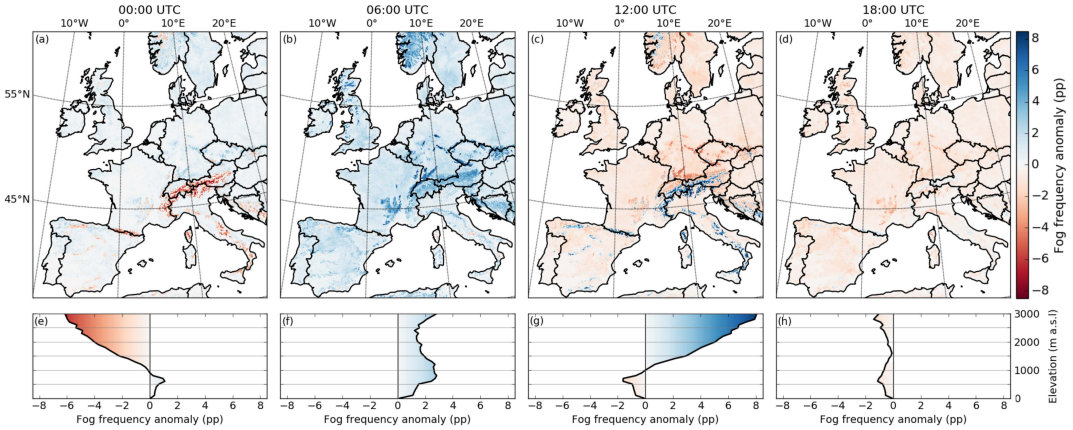

FIGURE 4.14: Fog frequency anomalies of (a) 00:00 UTC, (b) 06:00 UTC, (c) 12:00 UTC and (d) 18:00 UTC averaged over the complete period from 2006 to 2015. Anomaly values are given in in percentage points. Blue colors indicate positive values, red colors indicate negative values. Panels (e–h) give the corresponding averaged fog frequency anomalies for the different elevation levels.

the development of large fog patches within the whole valley. Another peculiarity of winter fog distribution is that the only negative anomaly can be found in the very high-altitude areas above 2200 m a.s.l. in the Alps and in the Pyrenees. This is probably due to the fact that these regions often rise above the FLS layers that are capped by the temperature inversion.

In spring, the fog anomaly in the lowlands shifts into the negative and in the higher regions above 900 m a.s.l. into the positive. This trend can be observed more strongly in the North than in the South, where also in spring higher values are reached than in the long-term average. The reason for this is the increasing condensation level while LST is still abundant. As a result, lowlands and lower mountain ranges are often below the condensation level, while the high mountain regions are covered in ground touching clouds more often.

This trend comes to an extreme during summer. Now, there is a strong overall negative fog anomaly and the areas that showed significantly more fog in winter show much less fog than in the long-term average. Even the high mountainous regions above 2000 m show a slight negative anomaly. High summer temperatures prevent the formation of deep lying stratus fields during anticyclonic conditions. As a result, neither the lowlands nor the slopes of the mountain ranges are covered in fog during these conditions. Even the highest regions are more frequently fog-free as only the fast moving clouds during cyclonic conditions are causing temporal low visibility situations here.

This trend is reversed again during autumn due to the general decrease of the condensation level. The lowlands start to be covered with fog more often again, while areas above 1300 m show a negative anomaly.

In summary, it can be concluded that the hourly and seasonal fluctuations in the spatial distribution of fog frequencies are adequately reflected by the product.

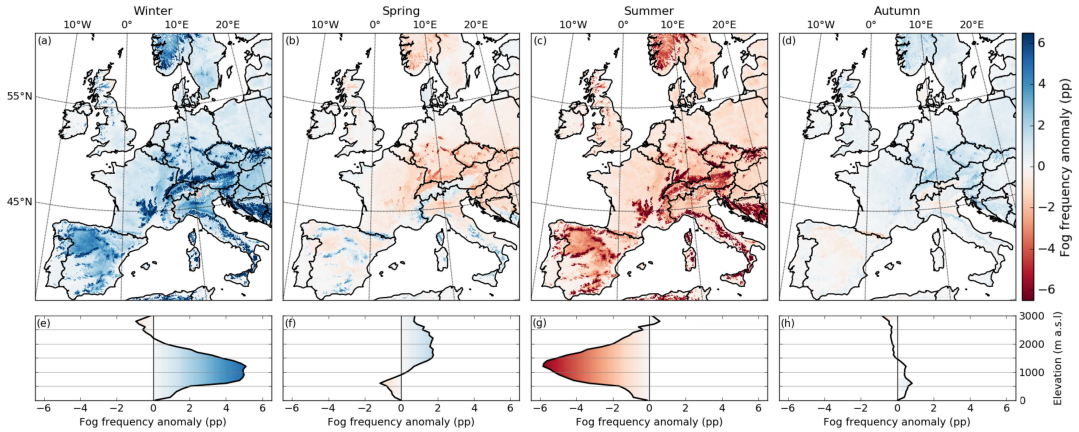

FIGURE 4.15: Fog frequency anomalies of the four seasons averaged over the complete period from 2006 to 2015. Anomaly values are given in in percentage points. Blue colors indicate positive values, red colors indicate negative values. (a) Winter: December, January, February; (b) Spring: March, April, May; (c) Summer: June, July, August; and (d) Autumn: September, October, November. Panels (e–h) give the corresponding averaged fog anomalies for the different elevation levels.

4.5 Conclusion

The hybrid fog retrieval approach presented in this paper is able to derive spatially explicit and temporally highly resoluted fog distribution data for Europe. Due to the incorporation of satellite and ground truth station data as well as additional terrain information it is able to capture the strong spatiotemporal dynamics of fog formation and dissipation sufficiently. The results of an extensive cross-validation procedure show that the generated fog product is capable of reflecting the diurnal and annual fluctuations in fog frequency very well.

The generated 10 year baseline fog climatology is the first of its kind. It provides valuable information on the fog frequency distribution over land on a 15 min interval and it reveals even small scale spatial patterns. The product shows that fog occurrence is highly dependent on the topography. High mountain ranges and uplands with a distinct relative elevation in comparison to their surroundings are characterized by high fog frequencies whereas lowlands and coast regions have much lower fog frequencies. During winter and night the strongest positive anomalies occur while in summer and during daytime the fog frequencies are much lower on average.

The product also shows, that amplitude and temporal lag of the fog frequency variation change with terrain elevation which can be attributed to the formation of different fog types. Similar results are found in recent studies in which radiation fog is associated with cold anticyclonic conditions in the lowlands and advection or cloud base lowering fog is linked with warm cyclonic conditions in elevated terrain (Akimoto and Kusaka, 2015; Belorid et al., 2014; Chen and Wang, 2015).

In order to develop a deeper understanding of the processes of fog dynamics, it is planned to investigate the fog product for different synoptic large-scale weather conditions and to determine allochthonous and autochthonous influence factors of fog formation and dissipation. Since both, METAR and SYNOP measurements, are available in quasi real-time, the developed algorithm can be used operationally in the future to continuously advance the product.

The product quality will be further improved by including additional station observations in different terrain types as well as more elevation levels and by using

pan-sharpening methods based on the HRV band in upcoming scheme adaptations. The algorithm can also be transferred without any major adjustments to the upcoming third generation of the Meteosat satellite series which will provide higher spatial, temporal and spectral resolutions. This should result in further improvements of the product accuracy. Monthly averages of the complete fog product data set are freely available for download at dx.doi.org/10.5678/LCRS/DAT.311.

Acknowledgements

This study was performed within the project "Ground fog detection and analysis with Machine Learning" (GFog-ML), generously funded by the German Research Foundation (DFG projects: BE1780/40-1 and TH1531/4-1). It is also part of the Satellite-based evaluation of weather conditioned LiDAR measurement downtimes project (LiMeS), generously funded by the Bundesministerium für Wirtschaft und Energie (BMWI project: 0324159D). The authors also thank EUMETSAT for providing the MSG satellite data used in this study and the DWD for providing the SYNOP station data.

Appendix

The following geostatistical texture features were used in this study. All features were calculated for 5x5 pixel surroundings with a lag distance of 1. The **Variogram** provides information about the correlation of the data with distance. It is defined as

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (v(x_i) - v(x_i + h))^2 \tag{4.1}$$

A slight variation to the Variogram is the **Madogram** which takes the absolute values of the differences:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} |v(x_i) - v(x_i + h)| \tag{4.2}$$

The **Rodogram** takes the square root of all differences:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \sqrt{|v(x_i) - v(x_i + h)|} \tag{4.3}$$

The **Cross Variogram** takes information from two different data inputs to account for cross correlation between two bands:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (v(x_i) - v(x_i + h))\,(w(x_i) - w(x_i + h)) \tag{4.4}$$

The **Pseudo Cross Variogram** calculates the differences of cross pairs between two data inputs directly:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} (v(x_i) - w(x_i + h))^2 \tag{4.5}$$

For all geostatistical texture features n(h) is the number of pairs for lag distance h. x_i and $x_i + h$ are the coordinates of pixel pairs that are separated by lag distance h for each pair i. v and w are the values of the pixels of different raster inputs at locations x_i and $x_i + h$ (Gloaguen et al., 2008).

The **Topographic Position Index** gives information about the relative elevation of a pixel in comparison to its neighbouring pixels. Positive values indicate a peak or a ridge, negative values indicate a valley. It is calculated with the following formula:

$$\mathrm{TPI} = v(x_i) - \frac{1}{n}\sum_{j=1}^{n} v(x_j) \tag{4.6}$$

where x_i is the index of the center pixel and x_j is the index of a neighbour pixel. v is the value of a pixel at a given index and n is the number of neighbouring pixels that are included in the calculation (Wilson and Gallant, 2000).

Distortions in the MSG bands due to the **Satellite Viewing Angle** were corrected for using the following formula:

$$c_i = b_i - t - s * sva_i + \bar{b} \tag{4.7}$$

where c_i is the value of pixel i of the corrected band, b_i is the value of the corresponding pixel of the original band, $\bar{b}$ is the average of all pixel values of the original band, sva_i is the SVA value of the corresponding pixel, s is the slope and t is the intercept of the linear regression that was determined before.

CBA prediction accuracy was evaluated using the **Mean Absolute Error**:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{Y}_i - Y_i\right| \tag{4.8}$$

where n is the number of predictions, $\hat{Y}_i$ is the predicted and Y_i is the observed value.

For the assessment of the confusion matrices that were produced in the comparison procedure of predicted and observed fog occurrences, the following **verification measures** were calculated:

$$\mathrm{PC} = \frac{a+d}{a+b+c+d} \tag{4.9}$$

$$\mathrm{POD} = \frac{a}{a+c} \tag{4.10}$$

$$\mathrm{FAR} = \frac{b}{a+b} \tag{4.11}$$

$$\mathrm{HSS} = \frac{2(ad-bc)}{(a+c)(c+d)+(a+b)(b+d)} \tag{4.12}$$

where a is the sum of all correctly predicted fog occurrences (True Positives) and b is the sum of all incorrectly predicted fog occurrences (False Positives). c is the sum of all False Negatives and d is the sum of all True Negatives accordingly.

Bibliography

Akimoto, Y. and Kusaka, H. (2015). A climatological study of fog in Japan based on event data. *Atmospheric Research*, 151:200–211.

Avotniece, Z., Klavins, M., and Lizuma, L. (2015). Fog climatology in Latvia. *Theoretical and Applied Climatology*, 122(1-2):97–109.

Bellman, R. E. (1961). *Adaptive Control Processes: A Guided Tour*. Princeton University Press.

Belorid, M., Lee, C. B., Kim, J. C., and Cheon, T. H. (2014). Distribution and long-term trends in various fog types over South Korea. *Theoretical and Applied Climatology*, 122(3-4):699–710.

Bendix, J. (1994). Fog climatology of the Po Valley. *Rivista di Meteorologia Aeronautica*, 54(3-4):25–36.

Bendix, J. (2002). A satellite-based climatology of fog and low-level stratus in Germany and adjacent areas. *Atmospheric Research*, 64(1-4):3–18.

Bendix, J. and Bachmann, M. (1991). Ein operationell einsetzbares Verfahren zur Nebelerkennung auf der Basis von AVHRR-Daten der NOAA-Satelliten. *Meteorologische Rundschau*, 43:169–178.

Bendix, J., Eugster, W., and Klemm, O. (2011). Fog – boon or bane? *Erdkunde*, 65(3):229–232.

Breiman, L. (1996). *Out-Of-Bag Estimation*. PhD thesis, University of California.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1):5–32.

Cermak, J. and Bendix, J. (2008). A novel approach to fog/low stratus detection using Meteosat 8 data. *Atmospheric Research*, 87(3-4):279–292.

Cermak, J. and Bendix, J. (2011). Detecting ground fog from space – a microphysics-based approach. *International Journal of Remote Sensing*, 32(12):3345–3371.

Cermak, J., Eastman, R. M., Bendix, J., and Warren, S. G. (2009). European climatology of fog and low stratus based on geostationary satellite observations. *Quarterly Journal of the Royal Meteorological Society*, 135(645):2125–2130.

Chen, H. and Wang, H. (2015). Haze Days in North China and the associated atmospheric circulations based on daily visibility data from 1960 to 2012. *Journal of Geophysical Research: Atmospheres*, 120(12):5895–5909.

Duda, D. P., Stephens, G. L., Stevens, B., and Cotton, W. R. (1996). Effects of Aerosol and Horizontal Inhomogeneity on the Broadband Albedo of Marine Stratus: Numerical Simulations. *Journal of the Atmospheric Sciences*, 53(24):3757–3769.

Egli, S., Maier, F., Bendix, J., and Thies, B. (2015). Vertical distribution of microphysical properties in radiation fogs — A case study. *Atmospheric Research*, 151:130–145.

Egli, S., Thies, B., Drönner, J., Cermak, J., and Bendix, J. (2017). A 10 year fog and low stratus climatology for Europe based on Meteosat Second Generation data. *Quarterly Journal of the Royal Meteorological Society*, 143(702):530–541.

EUMETSAT (2013). MSG Level 1.5 Image Data Format Description. Technical report, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt.

Finkensieper, S., van Meirink, J.-F., Zadelhoff, Gerd-Jan Hanschmann, T., Benas, N., Stengel, M., Fuchs, P., Hollmann, R., and Werscheck, M. (2016). CLAAS-2: CM SAF CLoud property dAtAset using SEVIRI - Edition 2.

Fuzzi, S., Facchini, M. C., Orsi, G., Lind, J. A., Wobrock, W., Kessel, M., Maser, R., Jaeschke, W., Enderle, K. H., Arends, B. G., Berner, A., Solly, I., Kruisz, C., Reischl, G., Pahl, S., Kaminski, U., Winkler, P., Ogren, J. A., Noone, K. J., Hallberg, A., Fierlinger-oberlinninger, H., Puxbaum, H., Marzorati, A., Hansson, H.-C., Wiedensohler, A., Svenningsson, I. B., Martinsson, B. G., Schell, D., and Georgii, H. W. (1992). The Po Valley Fog Experiment 1989. *Tellus B*, 44(5):448–468.

García-García, F. and Zarraluqui, V. (2008). A fog climatology for Mexico. *Erde*, 139(1-2):45–60.

Glickman, T. S. (2000). *Glossary of Meteorology (2nd ed.)*. American Meteorological Society, Boston, 2 edition.

Gloaguen, R., Wijaya, A., and Marpu, P. (2008). Geostatistical Texture Classification of Tropical Rainforest in Indonesia. In *Quality Aspects in Spatial Data Mining*, number 1, pages 199–210. CRC Press.

Gultepe, I., Tardif, R., Michaelides, S. C., Cermak, J., Bott, A., Bendix, J., Müller, M. D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., and Cober, S. G. (2007). Fog Research: A Review of Past Achievements and Future Perspectives. *Pure and Applied Geophysics*, 164(6-7):1121–1159.

Hamann, U., Walther, A., Baum, B., Bennartz, R., Bugliaro, L., Derrien, M., Francis, P. N., Heidinger, A., Joro, S., Kniffka, A., Le Gléau, H., Lockhoff, M., Lutz, H.-J., Meirink, J. F., Minnis, P., Palikonda, R., Roebeling, R., Thoss, A., Platnick, S., Watts, P., and Wind, G. (2014). Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms. *Atmospheric Measurement Techniques*, 7(9):2839–2867.

Henschel, J. R. and Seely, M. K. (2008). Ecophysiology of atmospheric moisture in the Namib Desert. *Atmospheric Research*, 87(3-4):362–368.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978.

Johnstone, J. a. and Dawson, T. E. (2010). Climatic context and ecological implications of summer fog decline in the coast redwood region. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10):4533–8.

Kawamoto, K. and Nakajima, T. (2001). A global determination of cloud microphysics with AVHRR remote sensing. *Journal of Climate*, 14(9):2054–2068.

Köhler, C., Steiner, A., Saint-Drenan, Y.-M., Ernst, D., Bergmann-Dick, A., Zirkelbach, M., Ben Bouallègue, Z., Metzinger, I., and Ritter, B. (2017). Critical weather situations for renewable energies – Part B: Low stratus risk for solar power. *Renewable Energy*, 101:794–803.

Kuhn, M. and Johnson, K. (2013). *Applied Predictive Modeling*. Springer New York, New York, NY.

Kühnlein, M., Appelhans, T., Thies, B., and Nauß, T. (2014). Precipitation Estimates from MSG SEVIRI Daytime, Nighttime, and Twilight Data with Random Forests. *Journal of Applied Meteorology and Climatology*, 53(11):2457–2480.

Lehnert, L. W., Thies, B., Trachte, K., Achilles, S., Osses, P., Baumann, K., Bendix, J., Schmidt, J., Samolov, E., Jung, P., Leinweber, P., Karsten, U., and Büdel, B. (2018). A Case Study on Fog/Low Stratus Occurrence at Las Lomitas, Atacama Desert (Chile) as a Water Source for Biological Soil Crusts. *Aerosol and Air Quality Research*, 18(1):254–269.

Mariani, L. (2009). Fog in the Po Valley: Some Meteo-Climatic Aspects. *Italian Journal of Agrometeorology*, 3(3):35–44.

Meyer, H., Kühnlein, M., Appelhans, T., and Nauss, T. (2015). Comparison of four machine learning algorithms for their applicability in satellite-based optical rainfall retrievals. *Atmospheric Research*, 169:424–433.

Musial, J., Hüsler, F., Sütterlin, M., Neuhaus, C., and Wunderle, S. (2014). Daytime Low Stratiform Cloud Detection on AVHRR Imagery. *Remote Sensing*, 6(6):5124–5150.

Nemery, B., Hoet, P. H. M., and Nemmar, A. (2001). The Meuse Valley fog of 1930: An air pollution disaster. *Lancet*, 357(9257):704–708.

Pedregosa, F., Gael Varoquaux, Alexandre Gramfort, Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition. *Journal of Machine Learning Research*, 12:2825–2830.

Pinto, R., Larrain, H., Cereceda, P., Lázaro, P., Osses, P., and S., S. R. (2001). Monitoring fog-vegetation communities at a fog-site in Alto Patache, South of Iquique, Northern Chile, during "El Niño" and "La Niña" events (1997-2000). In *Second International Conference on Fog and Fog Collection*, pages 293–296, Ottawa, Canada. International Development Research Center.

Scherrer, S. C. and Appenzeller, C. (2014). Fog and low stratus over the Swiss Plateau-a climatological study. *International Journal of Climatology*, 34(3):678–686.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A. (2002). An Introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological Society*, 83(7):977–992.

Schulz, H. M., Li, C.-F., Thies, B., Chang, S.-C., and Bendix, J. (2017). Mapping the montane cloud forest of Taiwan using 12 year MODIS-derived ground fog frequency data. *PLOS ONE*, 12(2):17.

Tardif, R. and Rasmussen, R. M. (2007). Event-Based Climatology and Typology of Fog in the New York City Region. *Journal of Applied Meteorology and Climatology*, 46(8):1141–1168.

Thies, B., Nauss, T., and Bendix, J. (2008). First results on a process-oriented rain area classification technique using Meteosat Second Generation SEVIRI nighttime data. *Adv. Geosci.*, 16:63–72.

Tjemkes, S. A., Van de Berg, L., and Schmetz, J. (1997). Warm Water Vapour Pixels over High Clouds as Observed by METEOSAT. *Contributions to atmospheric physics*, 70(1):15–22.

Vautard, R., Yiou, P., and van Oldenborgh, G. J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nature Geoscience*, 2(2):115–119.

Wendisch, M., Mertes, S., Heintzenberg, J., Wiedensohler, A., Schell, D., Wobrock, W., Frank, G., Martinsson, B. G., Fuzzi, S., Orsi, G., Kos, G., and Berner, A. (1998). Drop size distribution and LWC in Po Valley fog. *Contributions to atmospheric physics*, 71:87–100.

Wilson, J. P. and Gallant, J. C. (2000). *Terrain Analysis: Principles and Applications*. John Wiley & Sons, Ltd, New York, USA.

Chapter 5

A Spatially Explicit and Temporally Highly Resoluted Analysis of Variations in Fog Occurrence Over Europe

This chapter has been submitted to the *Quarterly Journal of the Royal Meteorological Society* and is currently in review under the title "A spatially explicit and temporally highly resoluted analysis of variations in fog occurrence over Europe."

Submitted: 07 December 2018

Abstract

Fog plays a major role in many ecological aspects and it influences human life in various ways. In this study, a temporally highly resoluted and spatially explicit analysis of variations in fog occurrence was conducted for Europe and links to general weather conditions were investigated. To this end, a high resolution fog product based on Meteosat Second Generation data was conducted. Characteristic fog distributions were identified by applying a Self Organizing Map approach onto the data set. It was found that the resulting fog patterns are primarily determined by terrain characteristics. Correlations between these patterns and the predominant general weather situations were computed and analyzed. The results show that the general weather situations can be categorized into three main groups, each responsible for the formation of a different group of fog patterns. Additionally, distinct regional differences could be identified in the diurnal and annual fog frequency cycles and the derived region-specific frequency variations were used to draw conclusions about the fog types prevailing in these regions.

5.1 Introduction

Fog is defined as a collection of suspended water droplets near the Earth's surface causing visibility to drop below 1000 m (Glickman, 2000). It has a strong effect on the environment in many aspects, e.g. as a water source for vegetation in otherwise arid regions (Henschel and Seely, 2008; Johnstone and Dawson, 2010; Lehnert et al., 2018; Pinto et al., 2001) and as a major regulator of entire ecosystems, such as in the montane cloud forests of Taiwan (Li et al., 2015; Schulz et al., 2017). Besides its ecological impacts, it also has a direct influence on human life. For instance, fog often obstructs traffic on roads and airports (Bendix et al., 2011; Dietz et al., 2018;

Kneringer et al., 2018) and destabilizes electricity grids by rapidly changing radiation conditions over photo-voltaic installations (Köhler et al., 2017). Fog also modifies the radiation balance of the atmosphere due to its high albedo. As a result, fog prevents air layers close to the ground from warming up by reducing the absorption of incoming solar radiation. This often leads to the persistence of inversions which, in turn, can lead to health consequences when toxic substances from the industry remain in the lower atmospheric layers in metropolitan areas for extended time spans (Nemery et al., 2001). Besides these negative effects, fog can also be beneficial to humans, especially to those living in arid regions. Here, fog water can be accumulated using collector systems when other water supplies are exhausted (Schemenauer et al., 1988; Schemenauer and Cereceda, 1991).

Due to all these factors, there is a growing demand for spatially accurate data on fog distribution and for a more precise understanding of the various meteorological drivers that govern the fog formation and dissipation processes. Especially in view of the anthropogenically induced climate change (IPCC, 2014) it can be assumed that the future will show changes in temporal and spatial fog frequency distributions, which most likely differ depending on fog type, geographic location and meteorological conditions. Latest research results have shown that in Europe and adjacent regions there has been a decrease in fog frequency at least since the 1960s (Avotniece et al., 2015; Gomez and Smith, 1984; Łupikasza and Niedźwiedź, 2016; Vautard et al., 2009; Veljović and Vujović, 2018). The main reasons for the decrease are seen in the improvement in air quality and rising air temperatures. Although these studies cover long time periods, they are only based on station data with few measurements per day. Therefore, they do not allow for spatially explicit statements and cannot adequately depict frequency shifts in the diurnal cycle. Furthermore, most of these studies focus on urban areas, where microclimatological phenomena have a different effect on fog occurrence than in rural areas. The only spatially explicit fog and low stratus baseline climatologies for Europe were presented in Cermak et al. (2009) and Egli et al. (2017). However, the focus of these studies was put on radiative cooling conditions only and no distinction was made between fog and low stratus which prevents these data sets from being used as a basis for an analysis of explicitly fog-influencing factors. In summary, there is a lack of spatially explicit information on fog distribution in Europe and the temporal coverage of existing analyses is insufficient. Due to the lack of spatially explicit information, the characteristics of fog distribution depending on the prevailing circulation patterns is consequently also not well understood.

To close these knowledge gaps and to be able to make future projections on the development of fog frequency in Europe, in this paper we give new insights into the spatial distribution and temporal variance in the occurrence of fog in Europe for the period from 2006 to 2015 using spatially explicit satellite data. Additionally, relations between fog distribution and general weather conditions are statistically analyzed. The study is based on a novel hybrid fog product that was computed by Egli et al. (2018). It derives fog information from a combination of geostationary satellite data and in situ station measurements. With the use of this product, we were able to carry out the first spatially explicit investigation of fog in Europe on a high temporal resolution with continuous temporal coverage and without the limitation to specific fog types.

The study was conducted as follows: First, the complete ten year data set was clustered into groups with specific but diverging spatial patterns on fog extent. For each group a representative fog pattern was identified. Then it was tested, whether the identified fog patterns could be associated with large-scale general weather situations and possible causes for the links found were investigated. Following the spatial pattern

analysis, annual and diurnal fog frequency cycles were examined in a time series analysis. Regions with different prevailing diurnal and annual fog frequency properties were identified and probable reasons for the distribution are formulated.

Section 5.2 introduces the data sets used. Chapter 5.3 gives an overview of the applied methodology. Results are presented and discussed in section 5.4. A final outlook is given in chapter 5.5.

5.2 Data

5.2.1 Fog Product

In order to investigate spatiotemporal distributions and shifts in fog frequency patterns in Europe, the fog product computed by Egli et al. (2018) was used. The product is based on a combination of Meteosat Second Generation (MSG) satellite data and Meteorological Aviation Routine Weather Reports (METAR) as well as synoptic weather observations (SYNOP) from the German Weather Service (DWD). The fog retrieval in this product was carried out by first predicting cloud base altitudes (CBA) for the entire domain using a random forest machine learning approach. In a second step, these CBA values were then adjusted using a dynamic threshold to identify areas with fog. The product is available for the time period from 2006 to 2015 at a temporal resolution of 15 min and can be accessed via DOI 10.5678/LCRS/DAT.311. Figure 5.1 shows two exemplary days with contrasting fog frequency patterns. On November 16, 2011 (a) vast areas of the lowlands in Central Europe were covered in radiation fog due to persistent anti-cyclonic conditions with increasing influx of cold air masses from Eastern Europe (DWD, 2012). In contrast, one month later, on December 16, 2011 (b) all mountain ranges in the domain showed high fog frequencies. This can be attributed to a strong western current in which the weather was dominated by fast traversing frontal systems (DWD, 2012) with stratus clouds causing low visibility conditions at the surface only in exposed terrain. In the following, fog frequencies will be referred to in raw percentages. In order to convert the raw percentages to fog hours per day ($\mathrm{h\,d^{-1}}$), they have to be multiplied by a factor of 0.24. For a conversion to fog days per year ($\mathrm{d\,a^{-1}}$), they have to be multiplied by 3.65.

5.2.2 General Weather Situations

Occurrences of principal fog frequency patterns were compared to the occurrences of general weather situations (GWS) over Europe. To this end, daily GWS data provided by the DWD which was derived using the classification scheme introduced by Hess and Brezowsky (1977) were obtained. The scheme divides all weather situations into 30 classes, each of which can be assigned to a zonal, meridional or a mixed circulation pattern. The procedure is based on a partitioning of the pressure fields over Europe and the North Atlantic which define the class membership of each weather situation. The GWL classification information is produced monthly by the DWD's forecasting and advisory centre and is made available on the DWD's website at https://www.dwd.de/DE/leistungen/grosswetterlage/grosswetterlage.html. Table 5.1 gives an overview of the general weather situations and their frequencies in the studied 10 year period.

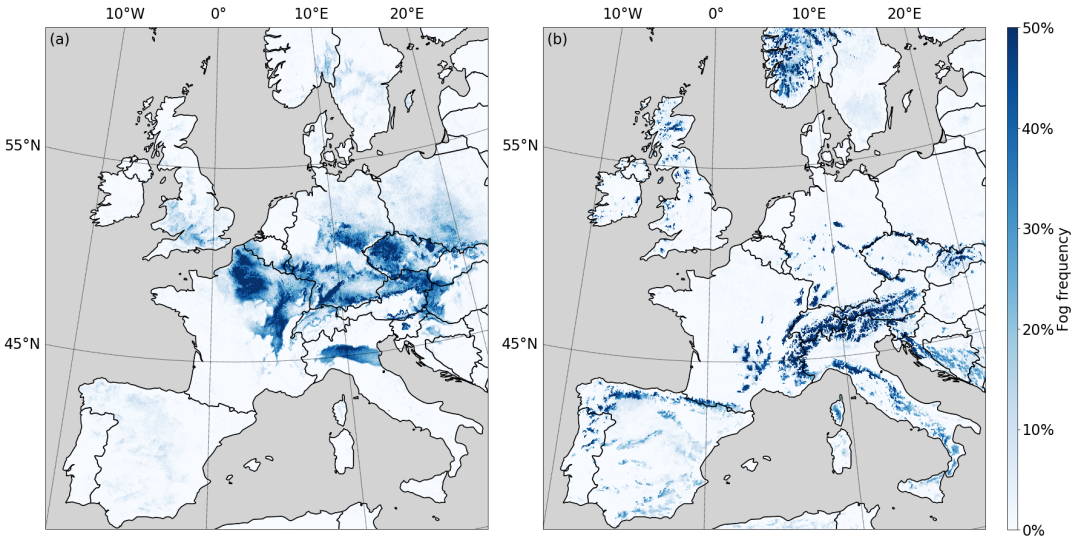

FIGURE 5.1: Fog frequency maps of (a) November 16, 2011 and (b) December 16, 2011. Fog frequencies are given in raw percentages.

GWS	Definition	Circ.	Min.	Max.	f (days)	f (%)
WA	Anticyclonic Westerly	zonal	Jan	Jul	188	5.15
WZ	Cyclonic Westerly	zonal	Apr	Aug	458	12.54
WS	South-Shifted Westerly	zonal	May	Feb	76	2.08
WW	Maritime Westerly (Block E. Europe)	zonal	Mar	Feb	74	2.03
SWA	Anticyclonic South-Westerly	mixed	May	Nov	98	2.68
SWZ	Cyclonic South-Westerly	mixed	Sep	Jun	280	7.67
NWA	Anticyclonic North-Westerly	mixed	May	Jun	94	2.57
NWZ	Cyclonic North-Westerly	mixed	Jul	Jan	183	5.01
HM	High over Central Europe	mixed	Sep	Jan	158	4.33
BM	Zonal Ridge across Central Europe	mixed	Dec	Mar	387	10.60
TM	Low (Cut-Off) over Central Europe	mixed	Jun	Aug	70	1.92
NA	Anticyclonic Northerly	meridional	Mar	Dec	27	0.74
NZ	Cyclonic Northerly	meridional	Mar	Dec	49	1.34
HNA	Icelandic High, Ridge C. Europe	meridional	Jul	May	71	1.94
HNZ	Icelandic High, Trough C. Europe	meridional	Nov	Apr	98	2.68
HB	High over the British Isles	meridional	Aug	Feb	106	2.90
TRM	Trough over Central Europe	meridional	Oct	Dec	330	9.04
NEA	Anticyclonic North-Easterly	meridional	Feb	May	35	0.96
NEZ	Cyclonic North-Easterly	meridional	Jun	Sep	66	1.81
HFA	Scandinavian High, Ridge C. Europe	meridional	Mar	Jan	76	2.08
HFZ	Scandinavian High, Trough C. Europe	meridional	Apr	Jan	54	1.48
HNFA	High Scand.-Iceland, Ridge C. Europe	meridional	May	Dec	17	0.47
HNFZ	High Scand.-Iceland, Trough C. Europe	meridional	Mar	Jun	88	2.41
SEA	Anticyclonic South-Easterly	meridional	Jun	Nov	74	2.03
SEZ	Cyclonic South-Easterly	meridional	Mar	Apr	46	1.26
SA	Anticyclonic Southerly	meridional	Mar	Nov	59	1.62
SZ	Cyclonic Southerly	meridional	Mar	May	46	1.26
TB	Low over the British Isles	meridional	Feb	Jul	77	2.11
TRW	Trough over Western Europe	meridional	Mar	Jul	228	6.24
UE	Transitional phase	none	Jul	Jan	39	1.07

TABLE 5.1: General Weather Situations after Hess and Brezowsky (1977). Definitions are taken from James (2007). All statistics refer to the period from 2006 to 2015. Frequency is given in absolute days and in percentage.

5.3 Methods

5.3.1 Fog Patterns

In order to get an overview of the spatial distribution of fog frequencies, characteristic fog situations were identified for the investigated 10-year period in a first step. This was done by assigning daily averaged frequency maps to specific fog patterns using a self-organizing map (SOM) algorithm. SOMs were first introduced by Kohonen (1982). They are based on single-layered artificial neural networks and are used to reduce the dimensionality of complex data structures. SOMs are widely applied in various scientific disciplines, including synoptic climatology and meteorology, where they are employed to derive patterns in atmospheric observations. (e.g. Hewitson and Crane, 2002; Liu and H., 2011; Sá et al., 2012; Tambouratzis and Tambouratzis, 2008). A SOM consists of a grid of neurons, each neuron representing its own class. In contrast to conventional neural networks in which the relative location of the neurons is irrelevant, neurons in SOMs do have a meaningful location which means that neighbouring neurons map similar properties, thus retaining the topology of the original data set. In our study this translates to adjacent neurons in the trained map representing similar fog patterns. This is particularly advantageous because it is impossible to define clear boundaries between different fog situations as the transitions between the individual situations are seamless. Another advantage of SOMs is that the size of the specified clusters can vary considerably. This makes the algorithm additionally favorable for our application as the frequencies of different fog situations are very heterogeneous.

The SOM network was instantiated with 30 nodes on a one-dimensional map in order to match the class count of the GWS classification. By this, the algorithm corresponds approximately to a constrained k-means clustering along the first principle component of the data. The clustering procedure was conducted using the PyMVPA package developed by Hanke et al. (2009). The implementation included in this library uses a squared Euclidean distance to determine the best matching node for each daily fog pattern and a Gaussian neighbourhood influence kernel. The resulting fog patterns of the SOM clustering algorithm are depicted in Figure 5.3. By using this method, every day was assigned to a specific fog pattern. This includes days without observed fog occurrences in the complete domain, which is just a special case, in short the "no fog" pattern.

5.3.2 Links to General Weather Situations

In order to check how individual fog patterns are favoured by certain weather conditions, the number of contemporaneous occurrences of a GWS together with a certain fog pattern was calculated separately for each GWS on a daily resolution. The formula for this calculation corresponds to the hit rate (HR):

$$\mathrm{HR} = \frac{\mathrm{d_f}}{\mathrm{d_f} + \mathrm{d_g}} \tag{5.1}$$

In this particular case, $\mathrm{d_f}$ stands for the number of days on which a certain fog pattern prevailed during the predominance of the respective GWS. $\mathrm{d_g}$ stands for the number of days on which this fog pattern did not prevail. In other words, the values calculated in this way indicate the probability for the occurrence of each fog pattern in the studied 10-year period during the persistence of a certain GWS. HR values were calculated for each possible combination between the GWSs and the fog patterns.

In order to check whether several GWSs could be combined into groups with similar fog pattern distributions, the HR matrix was additionally analyzed and sorted by means of a hierarchical clustering algorithm. In this study we used the fastcluster package developed by Müllner (2011, 2013). The linkage method included in this package is based on a recursive cluster merging algorithm that starts with a forest of clusters and merges two clusters with minimal distance in each step, finishing when only one cluster, the root cluster is left. The HR matrix is shown in Figure 5.5 together with the dendrogram resulting from the hierarchical clustering procedure. Since radiation and advection fogs in the lowland have the greatest impact on traffic and economy, specific GWSs favoring these conditions were analyzed in detail by calculating and inspecting fog frequency anomalies with respect to the 10-year average for these GWSs.

5.3.3 Diurnal and Annual Fog Frequency Cycles

In the third working step, the product of Egli et al. (2018) was analyzed for periodic oscillations in fog frequencies on a diurnal and on an annual time scale. For the diurnal analysis, fog frequencies were averaged over the complete 10 year data set separately for each 15 min interval of a day: 00:00 UTC, 00:15 UTC, ..., 23:45 UTC. For the examination of the annual periodicity, the data was averaged separately for each day of the year. This was also done over the complete 10 year period. The resulting average oscillation patterns are depicted in Figure 5.7. However, as the oscillation patterns strongly vary between different geographic locations, we analyzed fog frequency changes separately for each pixel of the data set. To this end, a sinus curve was fitted to the fog frequency distribution at each pixel. The sinus function was set up as follows:

$$\mathrm{f} = \sin\left(\frac{2\pi \mathrm{t}}{\mathrm{n}} + \mathrm{p}\right) * \mathrm{a} + \mathrm{o} \tag{5.2}$$

with f being the average fog frequency at time t. p is the phase, i.e. marks the location of the maximum/minimum in the diurnal/annual cycle. a is the amplitude which indicates how strong the diurnal/annual variation is and o is the vertical offset of the curve, thus the mean frequency over the complete period. n is the length of the period, in this case 96 quarter hours for the diurnal and 365/366 days for the annual period. The sinus function was fitted to the fog frequency data using the trust region reflective method developed by Branch et al. (1999) and implemented in the SciPy library (Jones et al., 2001).

Histograms of the resulting amplitude, phase and offset parameters are shown in Figure 5.8. In order to make the phase values easier to interpret, X-axis labels indicate the location of the maximum of the diurnal/annual cycle derived from the respective phase value using the following formula:

$$\mathrm{t_{max}} = \frac{\left(\left(\frac{\pi}{2} - \mathrm{p}\right) \bmod 2\pi\right) \mathrm{n}}{2\pi} \tag{5.3}$$

where $\mathrm{t_{max}}$ is the time of the maximum and n is the the length of the period, in this case 24 hours for the diurnal and 365/366 days for the annual period.

The three parameters were projected onto the red (offset), green (phase) and blue (amplitude) value ranges in the RGB-colorspace. Plotted as an RGB-composite on a map, they provide a comprehensive overview of the spatial variations of the fog frequency cycles with different colors representing different cycle properties. Since it is possible to identify distinct regions in these maps, the parameter combinations

were divided into separate groups in the next step. The k-means clustering algorithm implemented in the scikit-learn framework (Pedregosa et al., 2011) was employed for this purpose. In this way, it was possible to distinguish regions that differ in their fog occurrence and frequency patterns which could then be analysed for their geophysical properties, which in turn determine the meteorological and microphysical conditions that lead to fog formation and dissipation.

5.4 Results and Discussion

Depending on the field of application, different fog type classifications are made (cf. Byers, 1959; Cotton et al., 2011; Pérez-Díaz et al., 2017; Willett, 1928). Essentially relevant for Europe, however, are the following fog types: Radiation and advection fogs, orographic and mountain fogs as well as frontal fogs (see Figure 5.2). Radiation and advection fogs mainly occur in the lower flatland regions of Europe. Radiation fog forms due to nocturnal radiative cooling of air until the temperature falls below the dew point. When solar induced heat conduction from the ground and subsequent convective mixing cause the inversion to rise after sunrise, the fog layer is lifted or it dissipates completely. Advection fog forms as a consequence of vertical mixing when warm and humid air masses move over cold surfaces (Pérez-Díaz et al., 2017). As it is generally thicker than radiation fog and as it only dissipates when the wind pattern changes, it can often persist for multiple days. Orographic fog occurs as a result of adiabatic cooling due to orographically induced uplift of advected air masses (Bruijnzeel et al., 2005). It can be observed at the slopes along the windward side of higher mountain ranges when adiabatic cooling results in saturation and subsequently in condensation. Mountain fog, in contrast, is caused by the transport of clouds into mountainous regions causing low visibilities in exposed parts of the terrain (Eugster, 2008; Prettenthaler et al., 2010). This type of fog is a special type in so far as the condensation process has already taken place at another location and the air masses are only relocated afterwards. The formation of frontal fogs can be ascribed to an addition of water vapor to low lying air layers. It occurs at warm and cold fronts when precipitation from upper air layers falls into underlying near-saturated air masses and causes the water vapor saturation pressure to be exceeded in these layers (Cotton et al., 2011). It is therefore not coupled to the terrain properties but to the prevailing large scale weather situation. These fog types and their respective sub-classes are referred to in the following sections.

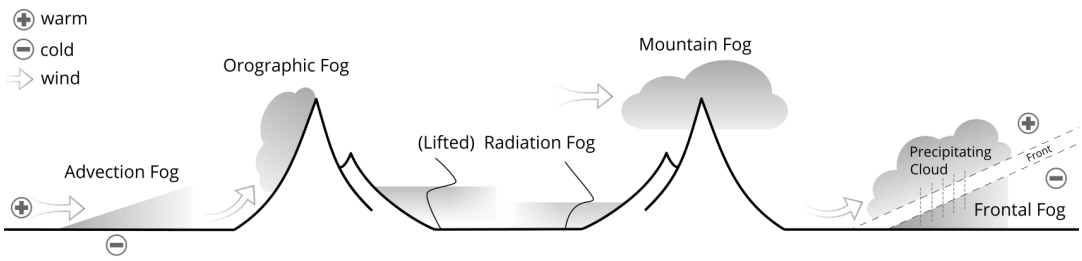

FIGURE 5.2: Schematic view of the fog type definitions used in this study. Arrows show wind directions. Plus and minus signs mark warm and cold air masses or surfaces. Black lines denote temperature profiles.

5.4.1 Fog Patterns

The SOM clustering procedure resulted in the 30 representative fog patterns depicted in Figure 5.3. The high variability in the number of days that make up the individual clusters (ranging from 59 in cluster 7 to 265 in cluster 3) shows that few fog patterns cover most of the time while many other fog patterns occur only rarely. It is noticeable that the differences between the different clusters are smooth: From low frequency fog patterns at the beginning to high frequency patterns around cluster 15 back to low frequency patterns in the end. Nevertheless, a cluster count of 30 was used for this analysis, as a clear distinction of individual fog patterns from each other is arbitrary and because rare but distinct fog patterns would be lost if they were combined into fewer clusters. In addition, the count of 30 corresponds to the count of GWSs examined in the subsequent steps and thus makes a direct comparison with the GWSs possible.

It is also noticeable that changes between the clusters are strongly oriented along the terrain. The reason for this is that the SOM algorithm forms clusters along the first principle component (the axis with the strongest variance). This axis is strongly influenced by the terrain because, besides general cloud abundance, terrain has a very strong explanatory content regarding fog occurrence (= ground touching clouds). Looking at Figure 5.4, it is also noticeable that certain cluster groups occur at different frequencies in different seasons. In detail, the fog patterns show the following properties:

Starting with cluster 1 in the upper left of Figure 5.3, the clusters generally contain little to no fog with minimal frequencies in clusters 3 and 4. Figure 5.4 shows that these non-fog patterns occur mainly in summer but also in the transitional seasons when the atmospheric conditions for cloud and fog formation due to radiative cooling are generally poorer. Next, high-frequency fog patterns in the low-lying regions of Europe can be seen around Cluster 8. The highest frequencies occur in the Po valley, in the Alpine foothills and in the Central European river valleys (e.g. in the upper Rhine valley and the Rhone valley). These fog patterns occur most frequently during autumn and winter when atmospheric conditions favour the formation of radiation fog (see Figure 5.4). Clusters 7 and 9 show a division into north-east (7) with higher frequencies in the lowlands of Eastern Europe, especially in Austria, the Czech Republic as well as Poland and south-west (9) with a frequency shift towards Southern France, Italy and Spain. Starting from cluster 10, the maximum frequencies increasingly shift into higher altitudes. Especially the lower mountain ranges, e.g. the Massif Central, the Vosges, the Black Forest, the Harz, the Swabian Mountains, the Bavarian Forest and the Tatra Mountains but also the foothills of the high mountains around the Alps and the Pyrenees are fog-covered here. The fog genesis is likely still caused by radiation weather conditions with lifted inversion layers when these patterns are formed or they are a result of advective transportation of warmer air masses causing condensation at the frontal zone. This results in condensed water reducing visibility mainly in elevated terrain. Until cluster 16 there is a further shift to high altitudes with generally high frequencies in mid and high mountain ranges, with fluctuations in fog frequency being mapped on the Iberian peninsula between cluster 13 and 18. The reason for fog formation in these cases may no longer be due to radiation weather conditions but rather to uplift processes on mountain slopes and associated condensation processes resulting in orographic fog. This assumption is also supported by the fact that these fog patterns do not show a clear annual cycle (see Figure 5.4). Up to cluster 26, the main fog zone then shifts further into the highest elevation regions. Here it can be assumed that low visibility can be explained by frontal clouds that are

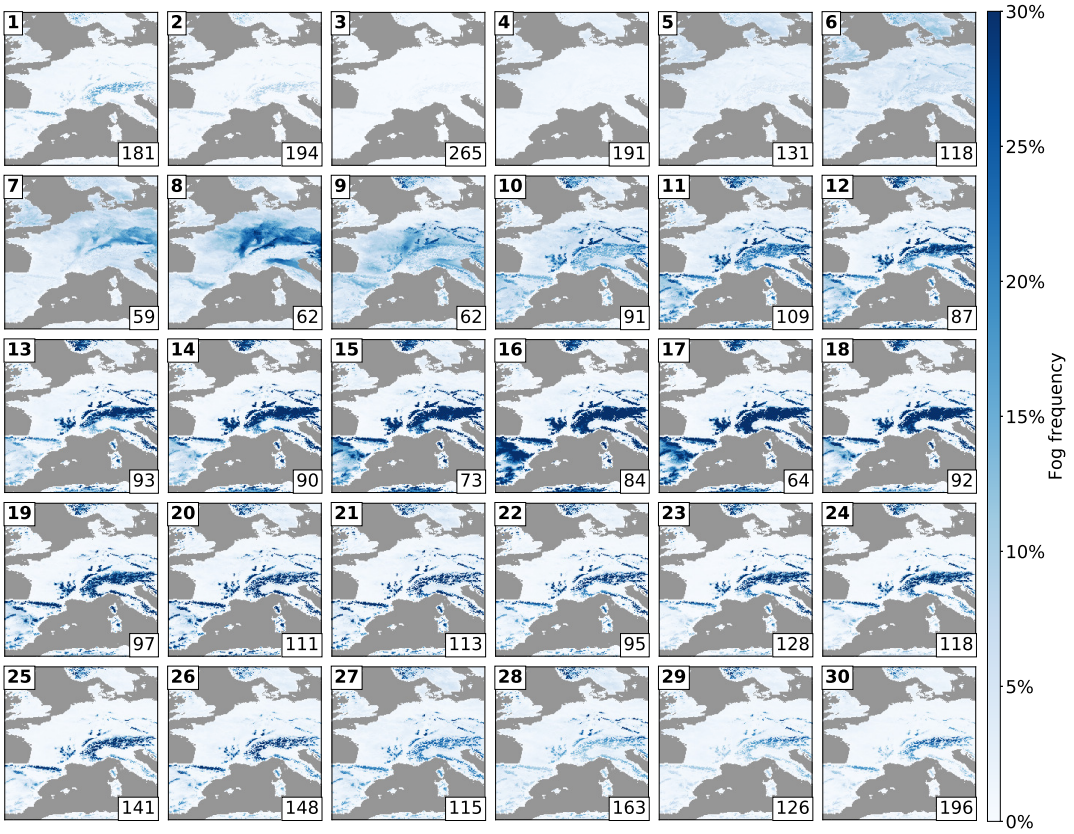

FIGURE 5.3: All 30 fog patterns derived using the SOM algorithm. The images show average fog frequency distributions of the central part of the studied domain in the original satellite projection for the respective fog pattern. Bold numbers in the top left show fog pattern IDs. Numbers in the bottom right show the number of days associated with the respective fog pattern.

coupled to fronts of traversing cyclone systems touching only the peaks of the mountains and thus causing mountain fog. After that, fog frequency generally decreases again towards cluster 30 which closes the circle to fog pattern 1 in the beginning.

5.4.2 Links to General Weather Situations

In order to establish a connection between the occurrence of certain fog patterns and the prevalence of certain weather conditions, the identified fog patterns were correlated with the GWS data published by the DWD as described in section 5.3 and were then clustered hierarchically. The result of this analysis is shown as a matrix in Figure 5.5. It is noticeable that there is a general dispersion in the matrix. A clear assignment of a GWS to a specific fog pattern is therefore not possible. The reason for this is that the GWSs are only defined by the atmospheric pressure patterns over Europe and the North Atlantic. The associated wind systems and distributions of temperature, humidity and aerosol content are thus only indirectly contained within the GWSs classification and can therefore differ significantly within one GWS. Thus, the conditions for fog formation are only given in certain cases within a GWS. It should also be mentioned that the classification of general weather situations according to Hess and Brezowsky (1977) is a subjective procedure and can therefore be handled differently by different meteorologists. Despite the dispersion, however, clear overlying

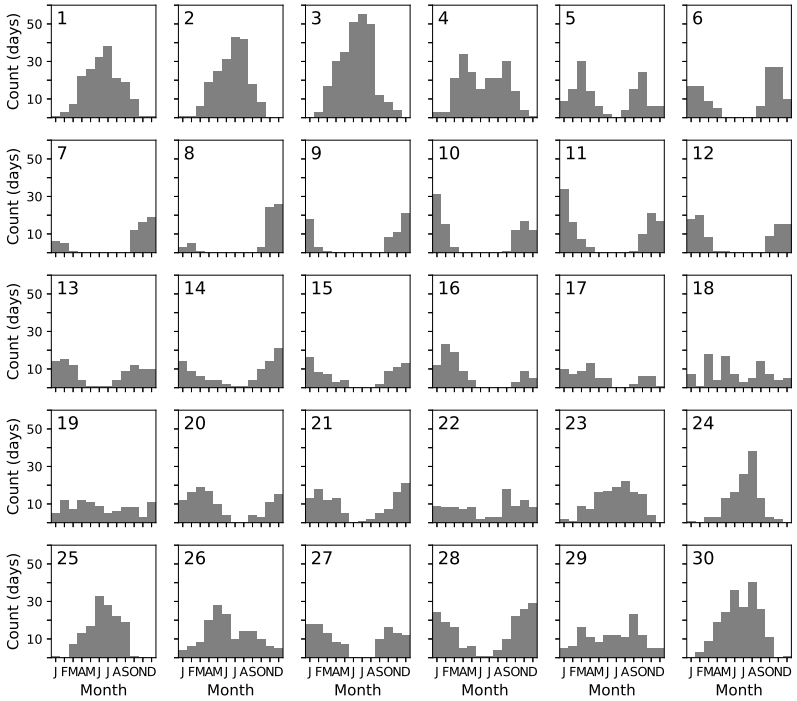

FIGURE 5.4: Histograms of annual distributions of the fog pattern occurrences that were identified with the SOM clustering algorithm. Data is accumulated over the complete 10 year period. Each bar represents one accumulated month, starting from January on the left.

patterns can be derived from the correlation matrix. The hierarchical clustering of the GWSs shows 3 main GWS groups, each of which can be linked to a different set of fog patterns (s. Figure 5.5 dendrogram and black horizontal lines).

The first group (anticyclonic) consists of the upper 13 GWSs, which are dominated by anticyclonic conditions (HB, WA, NWA, HFA, HNA, BM, HNFA). The most common fog patterns in this group are fog patterns that show very low fog frequencies throughout Europe (patterns 1-4 and 25-30). Only in the high mountain ranges one finds sporadic fog occurrence here. The reason for this distribution is probably to be found in the fact that these fog patterns occur most frequently in the summer season (see Figure 5.4). During this time, anticyclonic conditions generally cause low cloud and fog densities in contrast to the winter season when anticyclonic conditions favour the formation of low stratus clouds as well as radiation and advection fogs.

The second and smallest group (high pressure with southern influence) consists of the central four GWSs: SEA, SA, SWA and HM. All these GWSs are weather situations with anticyclonic high pressure conditions with inflow directions from southwest to southeast or calm wind conditions in the case of HM. The group shows a concentration around fog patterns 1 to 11. In addition to the "non-fog" patterns of 1 to 5, these are the radiation and advection fog patterns which exclusively cause fog in the lower flatland regions of Europe, mainly in winter but also in the transitional seasons (patterns 6 to 9). Other fog patterns with higher frequencies in elevated terrain do only very rarely occur during the prevalence of these weather conditions. The main reason for this distribution lies in the anticyclonality which favours the formation of fog in the winter season as a result of radiative cooling during calm periods or due to advection of moist air masses over cold land surfaces. The corresponding weather conditions with north winds (NEA, NA and NWA), on the other hand, show no increased

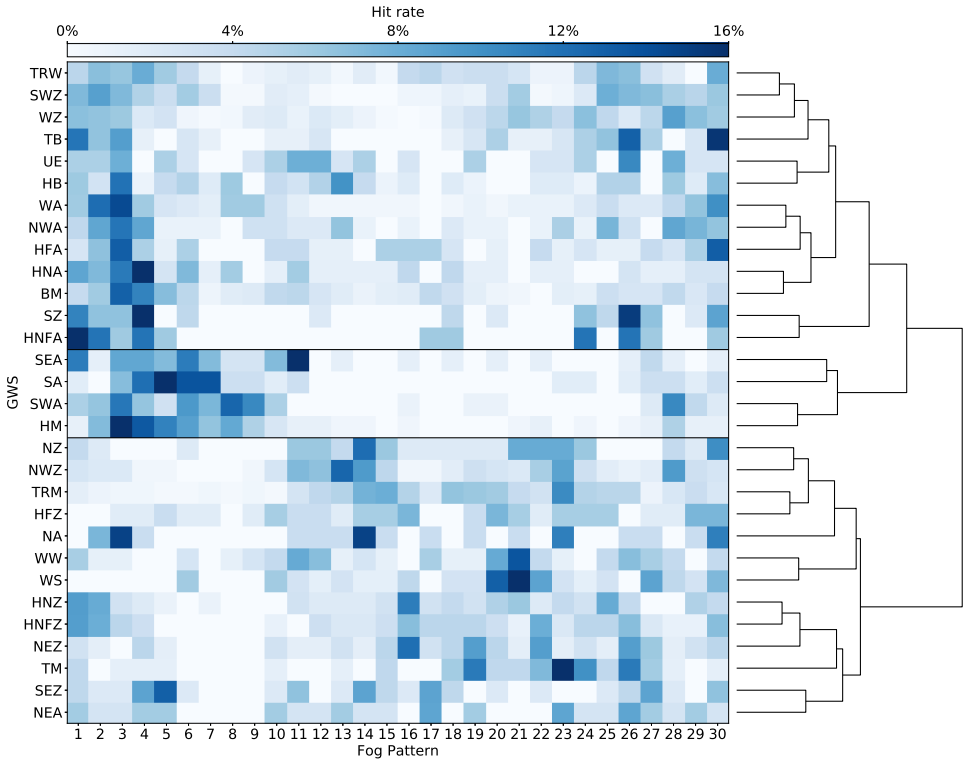

FIGURE 5.5: Hit rates for general weather situations and fog patterns. Each GWS line adds up to 100 %. The dendrogram on the right shows the resulting clustering from the hierarchical clustering procedure. The three main clusters are separated by black horizontal lines.

fog occurrence in the lowlands, despite their anticyclonic conditions. This is due to the fact that the air masses coming from the north are generally dryer and cooler than their counterparts from the south and thus no condensation sets in during the advection of these air masses over the Central European land masses. The basic fog-promoting properties of anticyclonic weather conditions in the lowlands, in contrast to cyclonic weather conditions, are also emphasized in field studies based on station data (e.g. Houssos et al., 2012; Łupikasza and Niedźwiedź, 2016; Wanner, 1979), even if local micrometeorological influences are found to be additional contributors to fog formation and dissipation processes here.

The third group (cyclonic) consists of the 13 lower GWSs, which are mainly affected by cyclonic conditions with the only exceptions of NA and NEA which occurred only very rarely during the investigated ten-year period (see Table 5.1). All GWSs of this group do not show a clear concentration around certain fog patterns but are rather randomly distributed over the matrix. However, the fog patterns between 6 and 9, i.e. those with high frequencies in the lowlands, are clearly missing here. This is due to the fact that cyclonic weather conditions never favour the formation of radiation fogs in the lowlands of Europe, independent of main wind direction and season. On the other hand, these weather conditions often lead to low visibilities in higher areas due to the formation of mountain fog and elevated stratus clouds coupled to traversing cyclone systems that touch the ground in exposed terrain.

Since it is the fog patterns of the second group in particular that have the greatest impact on traffic and economy, we take a closer look at the four GWSs in this group and their influence on the fog distribution in Europe. To this end, fog frequency anomaly maps were derived and analyzed for each of the four GWSs (see Figure 5.6).

During HM conditions (High over Central Europe), an extensive high pressure system lies over Central Europe. The Atlantic frontal zone runs in an anticyclonally curved arc usually north of 60° N. On the west and east flanks of the Central European high there are troughs over the Eastern Atlantic and over Russia. Sometimes a meridional high pressure zone forms. Air pressure gradients are often weak resulting in low wind speeds, low cloud cover and strongly pronounced temperature inversions throughout Central Europe (Hess and Brezowsky, 1977). The result is a significant increase in radiation fog occurrence in the lowlands and in the river valleys of Central Europe when low wind speeds and radiative cooling favor fog formation especially during winter and in the transitional seasons. This leads to a strong increase in fog formation in the lower flatlands everywhere in Central Europe and especially in the Po plain, the Saône and the upper Rhine valley. On the other hand, all mountain ranges show strong negative anomalies as the anticyclonality causes cloud-free conditions in the upper atmospheric layers most of the time. However, it must also be noted that by far the most common fog pattern that occurs under HM conditions is pattern no. 3, which shows minimal frequencies in the entire domain. This is due to the fact that despite high pressure conditions, radiation fog can only occur in those cases when also temperature, humidity and turbulence conditions allow its formation.

SWA conditions (Anticyclonic South-Westerly) cause a high pressure zone over southern Europe and western Russia and a low pressure system over the central North Atlantic and the western North Sea. Between these pressure fields, a frontal zone stretches from southwest to northeast, from the sea area southwest of Ireland to the Baltic States (Hess and Brezowsky, 1977). In comparison to HM conditions, this leads to an increase in fog frequencies mainly in Central and Eastern Europe but not in Great Britain and along the Atlantic coastlines. This can be attributed to individual perturbations moving northeastward along the frontal zone which only affect Great Britain and the western European coastal areas, whereas most regions of Central and Eastern Europe remain under anticyclonic influence. Another peculiarity is the strong increase in fog frequency over the Swiss Plateau. This is in good agreement with the findings of Troxler and Wanner (1991) and Wanner (1979), who were able to identify HM and SWA conditions to be among the major fog regulators in the lowlands of this region.

SA (Anticyclonic Southerly) is characterized by an extensive blocking high pressure area over Eastern Europe. Low air pressure prevails over the eastern Atlantic and parts of Western Europe. The Atlantic frontal zone runs from the Azores to southwest Europe, where it turns north. Individual disturbances only influence southwestern and western Europe (Hess and Brezowsky, 1977). This leads to high fog frequencies in Central Europe and England due to radiative cooling whereas the southern regions of Spain, France and Italy stay fog free.

During SEA prevalence (Anticyclonic South-Easterly) a high pressure ridge stretches from southeast Europe over the Baltic Sea and southern Scandinavia to the North Sea, sometimes as far as Iceland. Over the southeastern Atlantic lies a strong central depression. Peripheral disturbances of this depression essentially only affect Western Europe and occasionally influence Central Europe (Hess and Brezowsky, 1977). The atmospheric pressure distribution creates a southeastern inflow direction in the lower atmospheric levels which transports warm and humid air masses from the Mediterranean to Central Europe. During winter, cold air near the ground is here turbulently dispersed by the advected warm air layers. Because this results in a cooling of the warm and humid air masses, condensation sets in. This results in large low laying stratus patches, often covering wide areas of Central Europe. In contrast to HM, here we are not dealing with pronounced inversion situations that reach to the

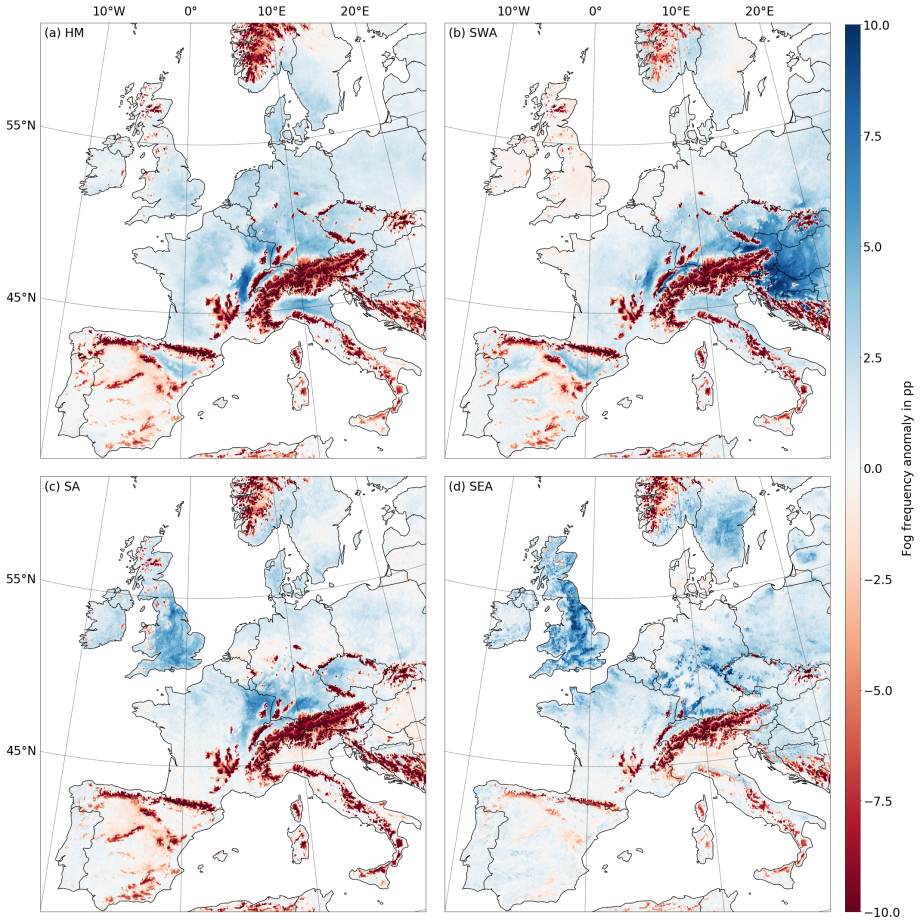

FIGURE 5.6: Anomalies of all GWSs related to lowland radiative and advective fog occurrence. Anomalies are calculated as deviations from the 10 year average and are given in percentage points (pp). For a conversion to hours per day ($\mathrm{h\,d^{-1}}$) they have to be multiplied by 0.24. (a) High over Central Europe, (b) Anticyclonic South-Westerly, (c) Anticyclonic Southerly, (d) Anticyclonic South-Easterly.

ground, so that during this GWS fog appears primarily on low to medium slopes of the lower mountain ranges in Central Europe. Thus, we are not dealing with ground touching radiation fog but with elevated low stratus patches. Higher terrain peaks reach over the stratus top and are consequently fog free. While positive anomalies can mainly be observed in Central Europe and in England, Spain, Southern France and the Po Valley show low to negative anomalies. These negative anomalies can probably be attributed to the influence of marginal disturbances from the depression above the southeast Atlantic.

In summary, it can be stated that different general weather conditions can, in principle, be linked to specific fog patterns on the basis of the investigated fog data set. Using the example of the lowland GWSs, it was also shown that spatial fog frequency distributions can be causally associated with the atmospheric pressure conditions of the respective GWS. However, a direct and unambiguous assignment of individual GWSs to certain fog patterns is not possible. The reason for this is that the atmospheric pressure field alone is only one explanatory factor for the formation and dissipation of fog. For a more precise assignment between the large scale meteorological conditions and the resulting fog distribution, further meteorological parameters would have to be taken into account in the analysis.

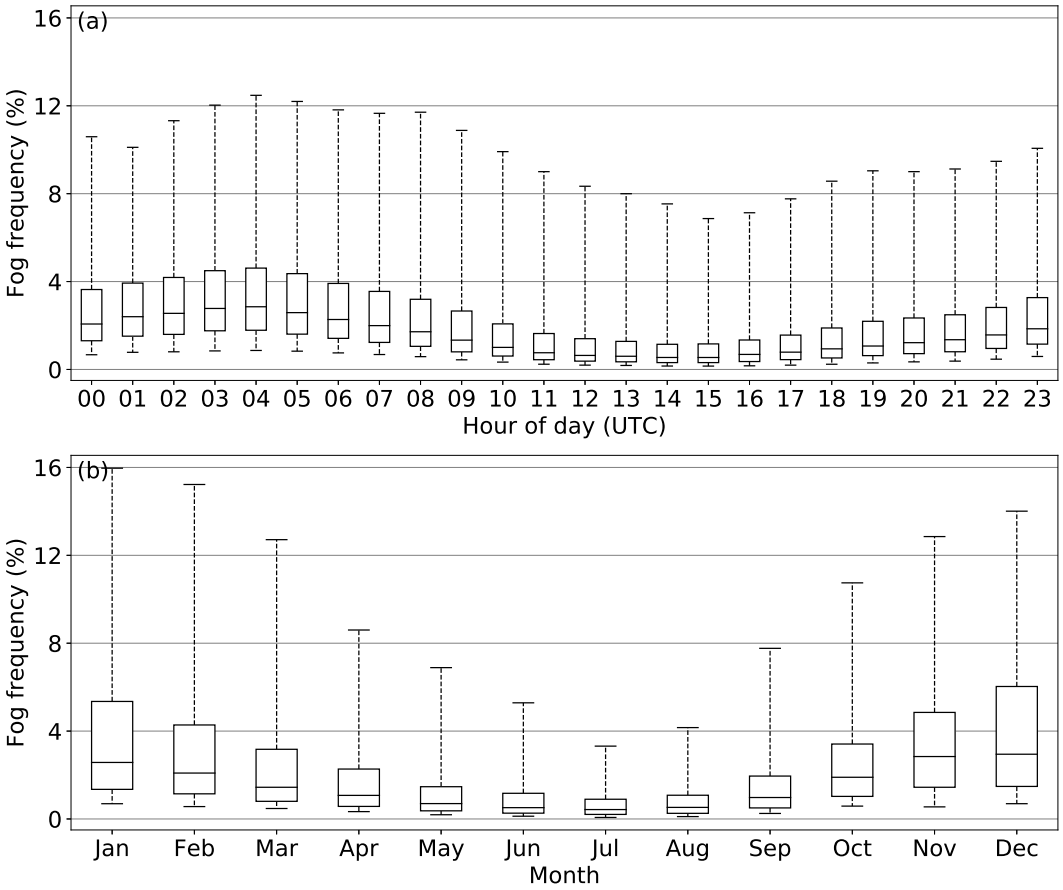

FIGURE 5.7: Average fog frequency changes over a day (a) and a year (b). Values are taken from each land pixel in the fog data set averaged over the complete ten year period. Whiskers mark 10 % and 90 % percentiles, boxes extend from the lower to the upper quartile of the fog frequency data with a line at the median.

5.4.3 Diurnal and Annual Fog Frequency Cycles

After examining the spatial distributions of different fog patterns and their relation to the large-scale weather conditions in Europe, we investigated the properties of the periodicity of fog frequencies in Europe. Looking at the fog frequency changes (averaged over all land pixels of the domain) in Figure 5.7 it gets obvious that two oscillation patterns overlay one another: There is a diurnal as well as an annual periodicity that can be discerned in the fog data. The highest values in the diurnal cycle are reached on average at 04:00 UTC in the morning (median: 2.8 %). The frequency values then decrease steadily until a minimum (median: 0.5 %) is reached at 15:00 UTC in the afternoon. Subsequently, the values rise continuously again. In the course of the year, a similar sinusoidal pattern can be seen, whereby the maximum is reached in December (median: 2.9 %) and the minimum in July (median: 0.4 %). In addition to the observed periodicity, it is also noticeable that the data exhibits strong variances within single time intervals. This confirms the assumption that different areas show strong differences in mean frequencies and periodic behavior. For this reason, it is necessary to look at diurnal and annual cycles separately for each region.

To this end, frequency cycles were investigated by fitting a sinus function to the diurnal and annual data variations at each single pixel as described in section 5.3.

The resulting parameter value distributions of the fitting procedure are depicted in Figure 5.8. Value ranges are clipped from the 1 % to the 99 % percentile. Amplitude values represent the strength of the diurnal and seasonal changes in fog frequency. Phase values are converted via Equation 5.3 such that they mark the time when the maximum in the respective cycle is reached. The offset indicates the mean value around which the data fluctuates, i.e. the mean fog frequency that is to be expected at the respective pixel. Here, it has to be emphasized, that a sinus curve can only be reasonably fitted to unskewed unimodal distributions. Multimodal and skewed distributions cannot be fitted very well which might result in insufficient fitting results in specific areas. The vast majority of fog frequency distributions in the investigated data set, however, follows a clear sinusoidal distribution when averaged over the ten year period. Thus, the resulting errors are expected to be small.

The diurnal amplitudes mainly range between 0.0 pp ($0.00\,\mathrm{h\,d^{-1}}$) and 5.5 pp ($1.32\,\mathrm{h\,d^{-1}}$) with a maximum around 1.1 pp ($0.26\,\mathrm{h\,d^{-1}}$). The distribution also shows that while most pixels show small changes during the course of the day between 0 pp ($0.00\,\mathrm{h\,d^{-1}}$) and 2 pp ($0.48\,\mathrm{h\,d^{-1}}$) there is only a small number of pixels with larger average amplitudes up to 5.5 pp ($1.32\,\mathrm{h\,d^{-1}}$). Looking at the distribution of the offsets one can assume that the high frequency of low amplitudes can be explained by the high frequency of low offsets, i.e. overall low fog occurrence. The strongly right-skewed distribution of the offsets shows that in the vast majority of regions the values fluctuate by about 1.4 pp ($0.34\,\mathrm{h\,d^{-1}}$) while there are very few isolated areas with average fog frequencies of up to over 26 pp ($6.2\,\mathrm{h\,d^{-1}}$). From the phase distribution it can be deduced that the maxima of the fog frequency responses are in most cases between 02:00 UTC and 05:00 UTC in the morning, but that there are also a few regions with maxima around noon.

The derived parameters of the annual oscillations basically show a similar distribution. However, the most frequent amplitude values here range below 0.5 pp ($1.83\,\mathrm{d\,a^{-1}}$), indicating a very weak annual cycle. On the other hand, a few pixels show amplitudes of over 10 pp ($36.6\,\mathrm{d\,a^{-1}}$), which shows that also regions with very strong annual cycles can be found. The phase values show that the most abundant fog frequency maxima lie in the winter half-year between December and January and the distribution of the offsets is, as in the case of daily oscillations, essentially limited to the range below 5 pp ($18.3\,\mathrm{d\,a^{-1}}$). When interpreting these values, it must be taken into account that they were calculated on the basis of average diurnal and annual variations. This has a reducing effect on the amplitude values, especially in the case of the daily variations. Here, data of 3597 days was averaged, whereas with the annual variations only 10 years were averaged.

A more precise analysis of the distribution of these parameter values can be achieved by looking at their spatial distribution in the domain. In order to get an overview of the spatial distribution, the three parameter value ranges were projected into the RGB color space and plotted as an RGB color composite as described in section 5.3. The results can be seen in Figure 5.9 (diurnal) and Figure 5.10 (annual) with subplots (a) showing the color composites and subplots (b) depicting the respective scatters in the RGB color cube, thus serving as "legends" with respect to subplots (a). Since multiple distinct regions can be identified in the RGB composites, the values were additionally divided into 6 groups with a k-means clustering algorithm as described in section 5.3. The resulting discrete classes are plotted in Figure 5.9 (c) and 5.10 (c) together with the average frequency curve of each class. Note that the ordinates of the blue clusters go up to 9 % while the ordinates of the green clusters reach 30 %. Although the boundaries between the clusters are arbitrary to a certain extent, one sees a reasonable division in the map.

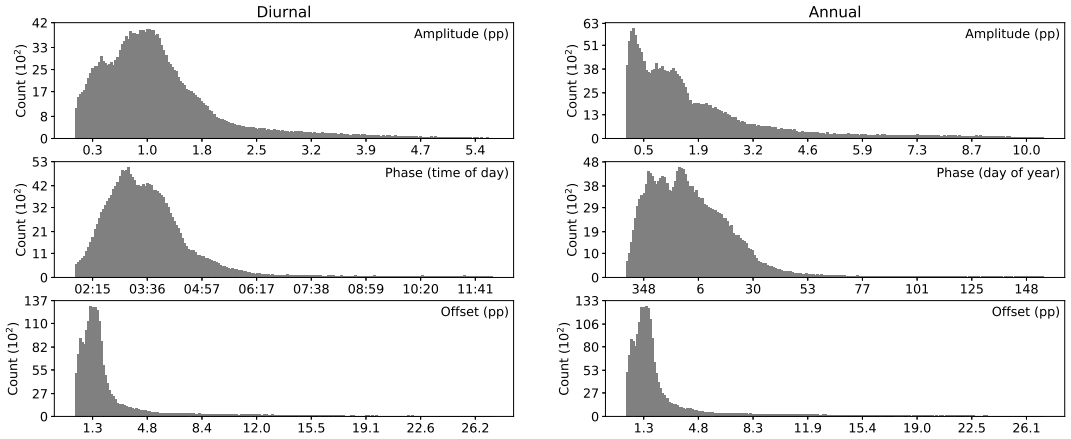

FIGURE 5.8: Histograms of amplitude, phase and offset of the sinus curve fitting results for diurnal (left) and annual (right) frequency cycles at each land pixel in the domain. Value ranges are clipped from the 1 % to the 99 % percentile.

The most striking feature of the diurnal map in Figure 5.9 (a) is that the majority of the low lying landmasses in Europe are colored yellow or orange which corresponds to the blue classes in Figure 5.9 (c). These are regions with minimum phase, amplitude and offset values which means we find overall low fog occurrences with little absolute changes here. It is noticeable that the phase values mainly range between 01:00 UTC and 04:00 UTC in these regions. This indicates that the majority of the occurring fog types can most likely be assigned to the radiation type which is consistent with the literature where radiation fogs are identified as the predominant type in the European lowlands (see Choularton et al., 1981; Fuzzi et al., 1992; Gultepe et al., 2007). It also explains the generally low frequency values since radiation fog occurrence is much rarer in total (especially if one considers the whole year) than fog events in the high mountains which are caused by orographic uplift processes or overlying cloud layers touching high and exposed terrain. Within the yellow area there is a shift in the phase values from 02:45 UTC to 04:00 UTC which manifests itself in a slight transition from yellow in the northeast to orange in the southwest and in the separation of the light-blue and the mid-blue cluster. This shift cannot be ascribed to geophysical differences in both regions. Rather, it reflects the temporal offset in the sun cycle (sunrise and sunset occur later in the west) and the associated shift in the diurnal fog life cycles.

The third fog type region (dark blue area in Figure 5.9 (c)) covers vast areas in Western Europe with major patches in the catchment area of the Duero in Spain, in the Saone valley in France, in the Swiss Plateau and in Southern Germany. Here, amplitudes of 1.7 pp (0.41 h d^{-1}) and offsets of 3.9 pp (0.94 h d^{-1}) are much higher than in the first two classes which results in a distinct increase in the oscillation intensity. In these areas, fog occurrence can also mainly be attributed to radiative cooling as indicated by the shape of the diurnal frequency curve with maxima still in the morning hours around 01:00 UTC to 04:00 UTC. In contrast to the first two fog type regions, geophysical properties in these areas obviously favour the formation of radiation fog more frequently. Depending on the specific region, these properties may differ. However, in most cases these are areas where the air exchange within the planetary boundary layer is vertically and horizontally severely restricted during windless high-pressure weather conditions in winter. Due to the surrounding topography, this leads to the formation of cold air collection areas, often in the main river valleys, which causes fog formation when temperatures drop below the condensation

level (Scherrer and Appenzeller, 2014; Troxler and Wanner, 1991; Wanner, 1979).

The fourth distinct fog type region is located at the cyan colored small areas in Figure 5.9 (a) which correspond to the light-green cluster in Figure 5.9 (c). These are areas with high offsets of 12.2 pp ($2.93\,\mathrm{h\,d^{-1}}$) and amplitudes of 3.5 pp ($0.84\,\mathrm{h\,d^{-1}}$) but low phase values (04:15 UTC). This means that we are confronted with areas with high fog frequencies and a clearly pronounced diurnal cycle with maxima in the morning hours. They are mainly located at the slopes of higher mountain ranges and in regions of the European low mountain ranges: The Massif Central, the Jura, the Vosges, the Black Forest, the Swabian, Bavarian, Saxonian and the Scandinavian mountain ranges as well as the Scottish Highlands. It can be assumed that the fog occurrences are mainly caused by advective and radiation fogs that are lifted from the ground to form low stratus patches so that they cause low visibilities only at the slopes of the low mountain ranges. The peaks and high mountains are not affected by these fog events because they rise above the stratus patches. These fog situations are often observed in countries like Switzerland, Germany and Austria where lowland areas are covered by low stratus, slope regions are covered in fog and higher mountain peaks are fog and cloud free (e.g. Prettenthaler et al., 2010; Scherrer and Appenzeller, 2014; Troxler and Wanner, 1991; Wanner, 1979).

The last main group is the dark blue area in the high mountains of the Pyrenees and the Alps in Figure 5.9 (a) which corresponds to the dark green area in Figure 5.9 (c). These are areas with strongly pronounced offsets of 24.8 pp ($5.95\,\mathrm{h\,d^{-1}}$) and amplitudes of 4.9 pp ($1.18\,\mathrm{h\,d^{-1}}$), i.e. high fog occurrence and strong diurnal cycles. In contrast to the areas of the low mountain ranges, however, here the maximum of the daily cycle lies around noon. This means that we are clearly no longer dealing with radiation fog types whose daily cycles are characterized by maxima before sunrise, but rather with orographic and mountain fog. Especially during summer, mountain fog primarily results from convective uplift processes and subsequent cloud formation, the intensity of which increases by noon. The strong heterogeneity of the diurnal cycles in the Alpine region (from yellow through cyan to blue) is consistent with the station observations made there (e.g. Prettenthaler et al., 2010). The data shows that the lowlands show smaller fog frequencies in total and maxima in the morning hours whereas higher elevated areas are dominated by higher fog frequencies and maxima around noon or in the afternoon. In the transition between the light-green and the dark-green cluster, we are dealing with regions in which amplitudes of 2.2 pp ($0.53\,\mathrm{h\,d^{-1}}$), offsets of 14.9 pp ($3.58\,\mathrm{h\,d^{-1}}$) and phase values (09:00 UTC) move towards the center of the value ranges. This results in an average shift of the phase values towards noon while average fog frequencies rise. This is probably because these areas are affected by all forms of fog genesis and show, in addition to individual high-reaching radiation and advection fogs, influence by mountain and orographic fog.

The parameter combinations derived from the annual oscillation patterns also show vast areas covered in yellow and orange (see Figure 5.10 (a)). Just as with the diurnal variations, this is a consequence of the fact that in most topographically flat regions there are generally significantly lower fog frequencies than in higher and exposed terrain. Here, the yellow/orange areas also correspond to the blue clusters of Figure 5.10 (c).

Starting with the light- and mid-blue clusters, one finds oneself in the coastal regions of the Mediterranean and the Atlantic with overall low offsets of 0.9 pp ($3.29\,\mathrm{d\,a^{-1}}$) and thus also low amplitudes of 0.3 pp ($1.10\,\mathrm{d\,a^{-1}}$). Phase values in the light-blue cluster are spread over a wide range, approximately from March to June with the average laying in the mid of May. The light- and mid-blue regions are more

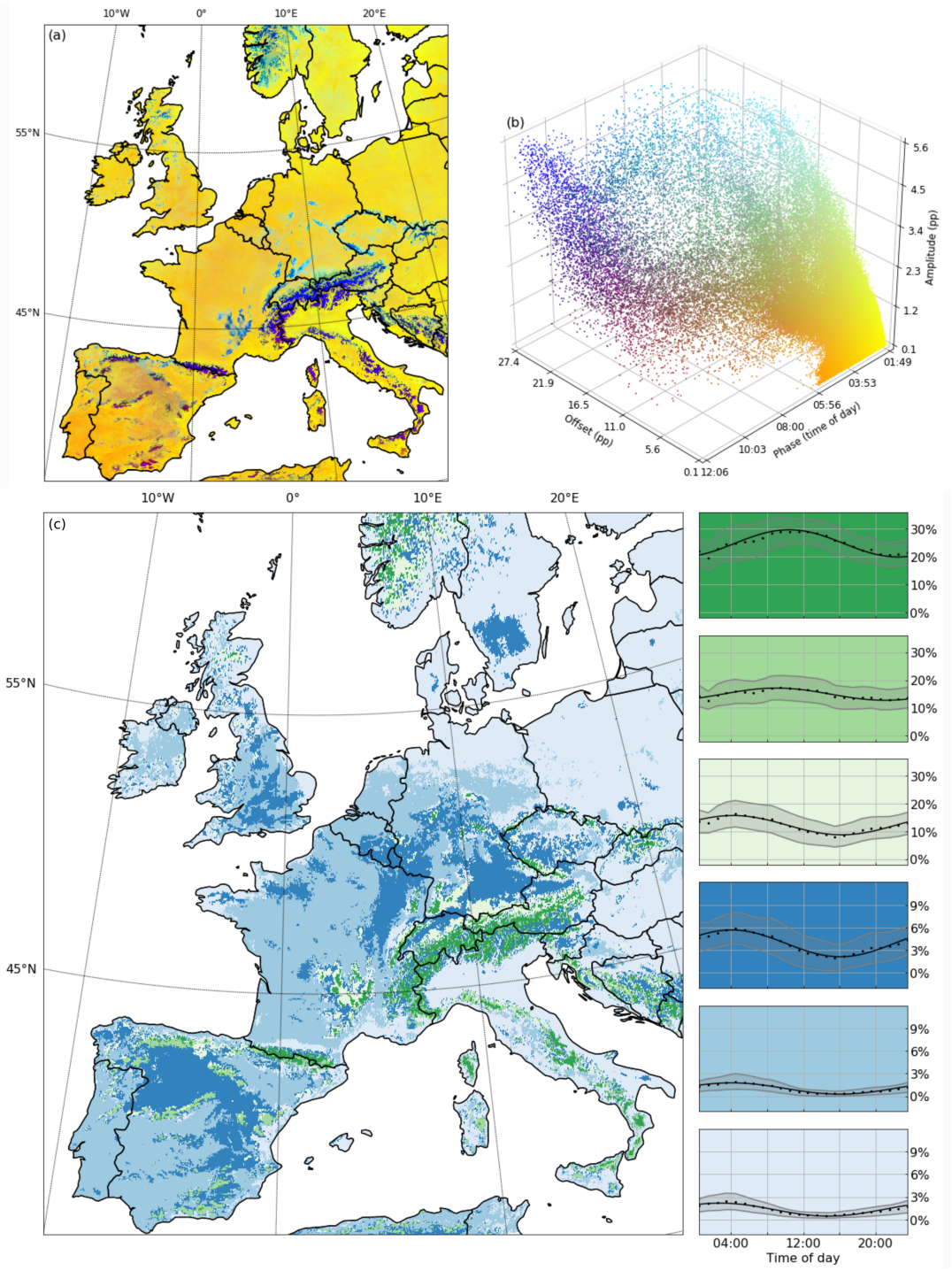

FIGURE 5.9: Parameter values derived from the sinusoidal fitting procedure applied to the **diurnal fog frequency variations**. (a) Amplitude, phase and offset plotted as RGB color composite. (b) 3D-Scatterplot of the parameter combinations colored according to the position within the RGB color cube. (c) Discrete classes of the parameter combinations derived using a k-means clustering method. Colored plots on the right show the average features of the respective class. The black line denotes the derived sinusoidal frequency curve, dots show the hourly-averaged fog frequencies of the original data and grey areas mark the standard deviation over all pixels within the respective cluster.

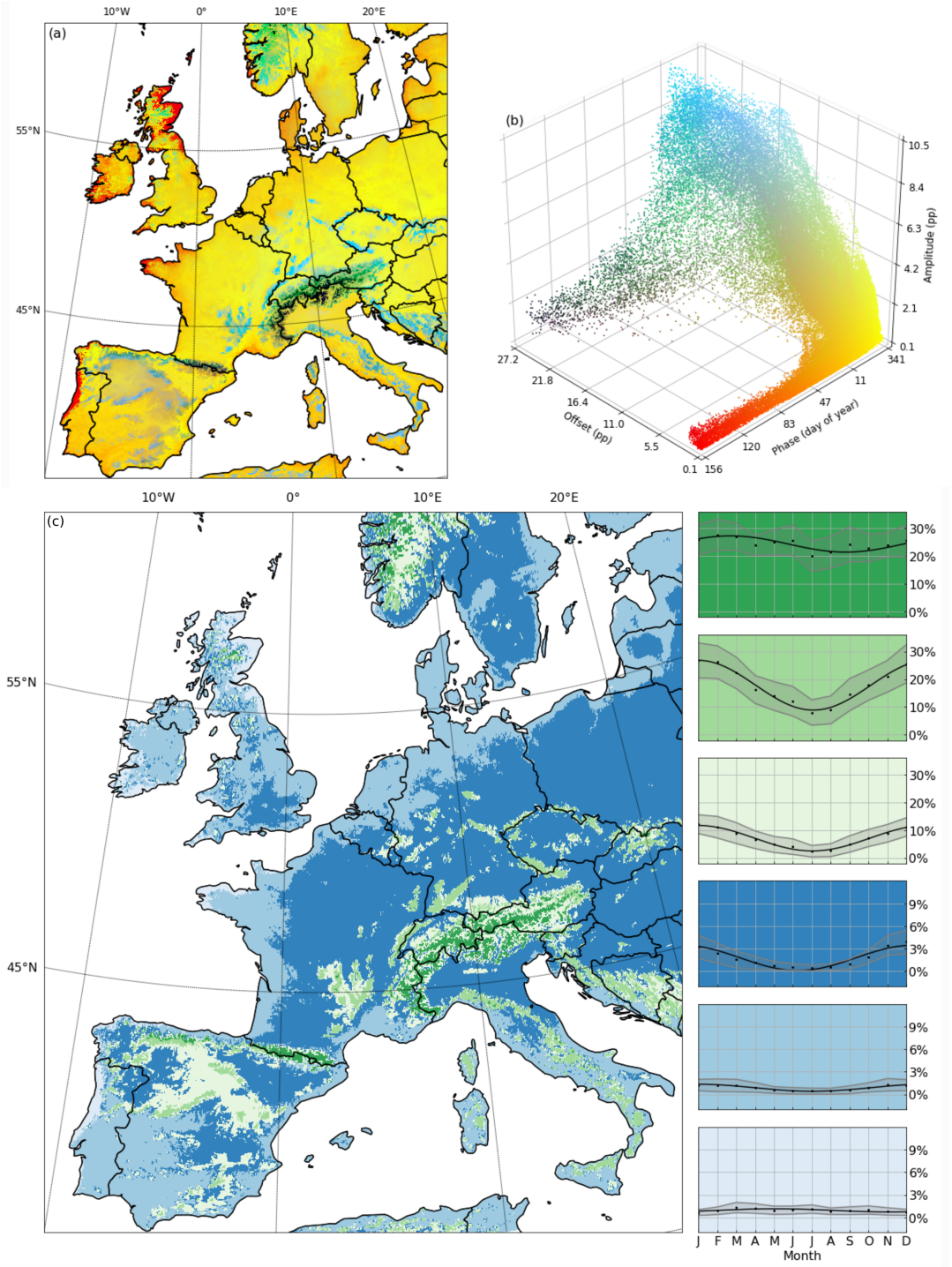

FIGURE 5.10: Parameter values derived from the sinusoidal fitting procedure applied to the **annual fog frequency variations**. (a) Amplitude, phase and offset plotted as RGB color composite. (b) 3D-Scatterplot of the parameter combinations colored according to the position within the RGB color cube. (c) Discrete classes of the parameter combinations derived using a k-means clustering method. Colored plots on the right show the average features of the respective class. The black line denotes the derived sinusoidal frequency curve, dots show the monthly-averaged fog frequencies of the original data and grey areas mark the standard deviation over all pixels within the respective cluster.

or less fog free throughout the year because neither topographic nor atmospheric conditions favor the formation of fog in these areas. Wind speeds during anticyclonic conditions are generally too high and thus stable temperature inversions cannot form which prevents the formation of radiation fog. Coastal fogs that form due to advection of moist and warm air masses over cold water surfaces are the only exception in these regions. They typically find their maximum occurrence during spring and in early summer when the atmosphere has already warmed up but the sea water is still cold (Findlater et al., 1989). This is in good agreement with the phase range of the light- and mid-blue clusters which is spread around spring and early summer months.

The dark-blue cluster region covers most of the remaining low-lying areas of Europe which are not directly situated at the coast lines but still show flat terrain in general. These areas are still characterized by relatively low offsets of 1.8 pp (6.59 d a^{-1}) on average. However, amplitudes strongly increased up to 1.7 pp (6.22 d a^{-1}) in these regions with maxima around mid of December. This is yet another indication that radiation fogs are the dominating type in these areas as they are known to form primarily during the cold season when radiative cooling at night allows for their formation.

Areas with a grayish tint in Figure 5.10 (a) corresponding to the light-green cluster in Figure 5.10 (c) are mainly found in the peripheral areas of the mountain ranges, e.g. the Meseta Central in Spain, the Massif Central in Southern France, the Scandinavian Mountains, the low mountain ranges in Germany, the Pyrenees and the Alps. Here, the annual variation is even more pronounced than in the dark-blue cluster with offsets of 7.1 pp (25.99 d a^{-1}) and amplitudes of 4.6 pp (16.84 d a^{-1}) while the curve maximum remains in winter around mid of January. This trend continues in the cyan-colored regions (mid-green cluster) that can be found in the peak areas of lower mountain ranges and along the slopes of the high mountains. While curve maxima still lie around mid of January, offsets of 17.8 pp (65.15 d a^{-1}) and amplitudes of 9.0 pp (32.94 d a^{-1}) significantly increase in these areas. The strongly pronounced annual cycle in these areas with maximum values in winter indicates that the fog genesis here too is mainly controlled by radiative cooling processes and advective fog formation. Fog frequency values in these areas are higher than in the adjacent lower terrain regions because they are still covered in fog even when the inversion layer is lifted from the ground in the lower areas. This situation occurs much more frequently than a true ground reaching inversion. In addition, these areas are often covered in low stratus clouds which are caused by advective circulation processes during unstable stratification. Lower terrain stays below these stratus patches and is thus fog-free.

The black-green colored areas in Figure 5.10 (a) which belong to the dark-green cluster in Figure 5.10 (c) differ considerably from the properties of the other clusters. This is especially true for the high altitude peak areas in the Alps and the Pyrenees. Extremely high fog frequencies occur throughout the year with an average offset of 24.4 pp (89.30 d a^{-1}). In contrast to the other clusters, however, only a very weakly pronounced annual cycle with amplitude values of 2.9 pp (2.56 d a^{-1}) and a maximum around the beginning of March can be seen here. This pattern cannot be explained by the predominance of radiation and advection fog types. Instead, low visibility conditions in these regions are mainly caused by higher stratus clouds coupled to passing frontal systems whose bases touch exposed mountainous areas. Since the conditions for these mountain fogs remain relatively constant throughout the year, no noteworthy annual frequency variation can be observed in these regions.

For an overview, all identified predominant fog types are summarized in Table 5.2 together with their spatio-temporal distributions as well as the atmospheric conditions and formation mechanisms controlling the respective fog type. However, this is just a basic list of the presumably dominant fog types in different areas which are of course

Fog type	Temp. occ.	Spat. occ.	Atmos. cond.	Form. mech.	C_d	C_a
Advection Fog	rarely, in winter, at night	Coast	Steady wind, warm air, cold surface	mixing		
Radiation Fog	sometimes, in winter, at night	Basins, valleys	Calm, low inversion, anticyclonal	radiative cooling		
Lifted Adv. & Rad. Fog	often, in winter, at night	Lower slopes	Light winds, high inversion	radiative cooling & mixing		
Orographic Fog	sometimes, all year, all day	Upper slopes	Strong winds, adiabatic uplift	adiabatic cooling		
Mountain Fog	very often, all year, at noon	High mountains	Stratiform or convective, cyclonal	relocation		

TABLE 5.2: All identified predominant fog types with main temporal and spatial occurrence, atmospheric conditions and basic formation mechanisms. C_d and C_a mark the color of the regions of the diurnal and annual fog frequency variation maps.

also influenced by multiple other fog types. As frontal fog is not that strongly tied to terrain characteristics, it cannot be assigned to a distinct region and thus does not appear in this list.

5.5 Conclusion

We were able to conduct the first spatially explicit investigation of fog patterns in Europe on the basis of a hybrid fog product developed by Egli et al. (2018). Groups of distinct fog patterns were identified and besides general cloud abundance, the terrain was identified to be a major regulator of fog distribution. A correlation analysis with general weather situations after Hess and Brezowsky (1977) was conducted which resulted in the identification of three main weather types that favor different combinations of fog patterns. In particular, anti-/cyclonality was identified as a good predictor for fog distribution. The most relevant fog patterns for ecology, economy and traffic with high frequencies in the lowlands could be assigned to anticyclonic weather conditions with a high over Central Europe with calm conditions (HM) and to high pressure weather situations with southwest to southeast inflow directions (SEA, SA, SWA). Additionally, diurnal and annual fog frequency cycles were investigated for the complete domain. That way, we were able to identify distinct regions with different cycle properties indicating different fog genesis types (radiative and advective fog vs. orographic and mountain fog as well as frontal fog).

However, the data set used only contains exclusively binary information on fog distribution derived from spectral information from the visible and thermal infrared range of the MSG satellite sensor. Direct information on the distribution of other meteorological variables such as temperature, wind, pressure, aerosol and turbulence fields were not included in this analysis. As the reasons for the formation and dissipation of fog are very complex and as they are dependent on all these variables, only assumptions about the main drivers of the identified fog types were made in the context of this work. An exact analysis of the underlying factors would have to be examined separately in individual studies for sample regions.

In order to gain further insights into the spatial differences in fog genesis, it is planned to include further meteorological and geophysical variables in the analysis, e.g. land cover information available through the CORINE Land Cover project as well

as air temperature and pressure fields from the ERA-5 assimilation data set produced by the Copernicus Climate Change Service or the regional COSMO-REA2 reanalysis product from the Hans-Ertel Center for Weather Research (Wahl et al., 2017). For a better delineation of fog patches and thus more precise information on fog frequency distribution, it is also planned to increase the resolution of the fog data set through the integration of the high resolution visible (HRV) band of MSG and by exploiting data from low earth orbiting satellite systems like the Moderate Resolution Imaging Spectroradiometer (MODIS) or the Advanced Very High Resolution Radiometer (AVHRR). In addition, we are currently working on the operationalization of the processing pipeline of the fog product in order to be able to continuously expand our analysis beyond the year 2015.

Acknowledgements

This study was performed within the project "Ground fog detection and analysis with Machine Learning" (GFog-ML), generously funded by the German Research Foundation (DFG projects: BE1780/40-1 and TH1531/4-1). It is also part of the Satellite-based evaluation of weather conditioned LiDAR measurement downtimes project (LiMeS), generously funded by the Bundesministerium für Wirtschaft und Energie (BMWI project: 0324159D). The authors also thank the DWD (German Weather Service) for providing the GWS data.

Bibliography

Avotniece, Z., Klavins, M., and Lizuma, L. (2015). Fog climatology in Latvia. *Theoretical and Applied Climatology*, 122(1-2):97–109.

Bendix, J., Eugster, W., and Klemm, O. (2011). Fog – boon or bane? *Erdkunde*, 65(3):229–232.

Branch, M. A., Coleman, T. F., and Li, Y. (1999). A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems. *SIAM Journal on Scientific Computing*, 21(1):1–23.

Bruijnzeel, S., Eugster, W., and Burkard, R. (2005). Fog as a Hydrologic Input. In *Encyclopedia of Hydrological Sciences*, pages 559–582. John Wiley & Sons, Ltd, Chichester, UK.

Byers, H. R. (1959). *General meteorology*. McGraw-Hill, New York, 3 edition.

Cermak, J., Eastman, R. M., Bendix, J., and Warren, S. G. (2009). European climatology of fog and low stratus based on geostationary satellite observations. *Quarterly Journal of the Royal Meteorological Society*, 135(645):2125–2130.

Choularton, T. W., Fullarton, G., Latham, J., Mill, C. S., Smith, M. H., and Stromberg, I. M. (1981). A field study of radiation fog in meppen, West Germany. *Quarterly Journal of the Royal Meteorological Society*, 107(452):381–394.

Cotton, W. R., Bryan, G., and van den Heever, S. C. (2011). Fogs and Stratocumulus Clouds. In *International Geophysics*, volume 99, pages 179–242.

Dietz, S. J., Kneringer, P., Mayr, G. J., and Zeileis, A. (2018). Forecasting Low-Visibility Procedure States with Tree-Based Statistical Methods. *Pure and Applied Geophysics*.

DWD (2012). Großwetterlage (Januar bis Dezember 2011). Technical report, Deutscher Wetterdienst, Offenbach.

Egli, S., Thies, B., and Bendix, J. (2018). A Hybrid Approach for Fog Retrieval Based on a Combination of Satellite and Ground Truth Data. *Remote Sensing*, 10(4):26.

Egli, S., Thies, B., Drönner, J., Cermak, J., and Bendix, J. (2017). A 10 year fog and low stratus climatology for Europe based on Meteosat Second Generation data. *Quarterly Journal of the Royal Meteorological Society*, 143(702):530–541.

Eugster, W. (2008). Fog Research. *Die Erde*, 139:1–10.

Findlater, J., Roach, W. T., and McHugh, B. C. (1989). The haar of north-east Scotland. *Quarterly Journal of the Royal Meteorological Society*, 115(487):581–608.

Fuzzi, S., Facchini, M. C., Orsi, G., Lind, J. A., Wobrock, W., Kessel, M., Maser, R., Jaeschke, W., Enderle, K. H., Arends, B. G., Berner, A., Solly, I., Kruisz, C., Reischl, G., Pahl, S., Kaminski, U., Winkler, P., Ogren, J. A., Noone, K. J., Hallberg, A., Fierlinger-oberlinninger, H., Puxbaum, H., Marzorati, A., Hansson, H.-C., Wiedensohler, A., Svenningsson, I. B., Martinsson, B. G., Schell, D., and Georgii, H. W. (1992). The Po Valley Fog Experiment 1989. *Tellus B*, 44(5):448–468.

Glickman, T. S. (2000). *Glossary of Meteorology (2nd ed.)*. American Meteorological Society, Boston, 2 edition.

Gomez, B. and Smith, C. G. (1984). Atmospheric Pollution and Fog Frequency in Oxford, 1926-1980. *Weather*, 39(12):379–384.

Gultepe, I., Tardif, R., Michaelides, S. C., Cermak, J., Bott, A., Bendix, J., Müller, M. D., Pagowski, M., Hansen, B., Ellrod, G., Jacobs, W., Toth, G., and Cober, S. G. (2007). Fog Research: A Review of Past Achievements and Future Perspectives. *Pure and Applied Geophysics*, 164(6-7):1121–1159.

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., and Pollmann, S. (2009). PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data. *Neuroinformatics*, 7(1):37–53.

Henschel, J. R. and Seely, M. K. (2008). Ecophysiology of atmospheric moisture in the Namib Desert. *Atmospheric Research*, 87(3-4):362–368.

Hess, P. and Brezowsky, H. (1977). *Katalog der Grosswetterlagen Europas 1881–1976, 3. verbesserte und ergänzte Aufl.* Deutscher Wetterdienst, Offenbach am Main, Germany.

Hewitson, B. and Crane, R. (2002). Self-organizing maps: applications to synoptic climatology. *Climate Research*, 22(1):13–26.

Houssos, E. E., Lolis, C. J., Gkikas, A., Hatzianastassiou, N., and Bartzokas, A. (2012). On the atmospheric circulation characteristics associated with fog in Ioannina, north-western Greece. *International Journal of Climatology*, 32(12):1847–1862.

IPCC (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge.

James, P. M. (2007). An objective classification method for Hess and Brezowsky Grosswetterlagen over Europe. *Theoretical and Applied Climatology*, 88(1-2):17–42.

Johnstone, J. a. and Dawson, T. E. (2010). Climatic context and ecological implications of summer fog decline in the coast redwood region. *Proceedings of the National Academy of Sciences of the United States of America*, 107(10):4533–8.

Jones, E., Oliphant, T., Peterson, P., and Others (2001). SciPy: Open source scientific tools for Python.

Kneringer, P., Dietz, S. J., Mayr, G. J., and Zeileis, A. (2018). Probabilistic Nowcasting of Low-Visibility Procedure States at Vienna International Airport During Cold Season. *Pure and Applied Geophysics*, (September).

Köhler, C., Steiner, A., Saint-Drenan, Y.-M., Ernst, D., Bergmann-Dick, A., Zirkelbach, M., Ben Bouallègue, Z., Metzinger, I., and Ritter, B. (2017). Critical weather situations for renewable energies – Part B: Low stratus risk for solar power. *Renewable Energy*, 101:794–803.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1):59–69.

Lehnert, L. W., Thies, B., Trachte, K., Achilles, S., Osses, P., Baumann, K., Bendix, J., Schmidt, J., Samolov, E., Jung, P., Leinweber, P., Karsten, U., and Büdel, B. (2018). A Case Study on Fog/Low Stratus Occurrence at Las Lomitas, Atacama Desert (Chile) as a Water Source for Biological Soil Crusts. *Aerosol and Air Quality Research*, 18(1):254–269.

Li, C.-F., Zelený, D., Chytrý, M., Chen, M.-Y., Chen, T.-Y., Chiou, C.-R., Hsia, Y.-J., Liu, H.-Y., Yang, S.-Z., Yeh, C.-L., Wang, J.-C., Yu, C.-F., Lai, Y.-J., Guo, K., and Hsieh, C.-F. (2015). Chamaecyparis montane cloud forest in Taiwan: ecology and vegetation classification. *Ecological Research*, 30(5):771–791.

Liu, Y. and H., R. (2011). A Review of Self-Organizing Map Applications in Meteorology and Oceanography. In *Self Organizing Maps - Applications and Novel Algorithm Design*, number January, pages 253–272. InTech.

Łupikasza, E. and Niedźwiedź, T. (2016). Synoptic climatology of fog in selected locations of southern Poland (1966–2015). *Bulletin of Geography. Physical Geography Series*, 11(1):5–15.

Müllner, D. (2011). *Modern hierarchical, agglomerative clustering algorithms*. Number 1973.

Müllner, D. (2013). fastcluster : Fast Hierarchical, Agglomerative Clustering Routines for R and Python. *Journal of Statistical Software*, 53(9).

Nemery, B., Hoet, P. H. M., and Nemmar, A. (2001). The Meuse Valley fog of 1930: An air pollution disaster. *Lancet*, 357(9257):704–708.

Pedregosa, F., Gael Varoquaux, Alexandre Gramfort, Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition. *Journal of Machine Learning Research*, 12:2825–2830.

Pérez-Díaz, J., Ivanov, O., Peshev, Z., Álvarez-Valenzuela, M., Valiente-Blanco, I., Evgenieva, T., Dreischuh, T., Gueorguiev, O., Todorov, P., and Vaseashta, A. (2017). Fogs: Physical Basis, Characteristic Properties, and Impacts on the Environment and Human Health. *Water*, 9(10):807.

Pinto, R., Larrain, H., Cereceda, P., Lázaro, P., Osses, P., and S., S. R. (2001). Monitoring fog-vegetation communities at a fog-site in Alto Patache, South of Iquique, Northern Chile, during "El Niño" and "La Niña" events (1997-2000). In *Second International Conference on Fog and Fog Collection*, pages 293–296, Ottawa, Canada. International Development Research Center.

Prettenthaler, F., Podesser, A., and Pilger, H. (2010). *Klimaatlas Steiermark*. Verlag der Österreichischen Akademie der Wissenschaften, Wien, 4 edition.

Sá, J. A., Rocha, B., Almeida, A., and Souza, J. R. (2012). Recurrent Self-Organizing Map for Severe Weather Patterns Recognition. In ElHefnawi, M., editor, *Recurrent Neural Networks and Soft Computing*, pages 151–174. InTech.

Schemenauer, R. S. and Cereceda, P. (1991). Fog-water Collection in Arid Coasta Locations. *Ambio*, 20(7):303–308.

Schemenauer, R. S., Fuenzalida, H., and Cereceda, P. (1988). A Neglected Water Resource: The Camanchaca of South America. *Bulletin of the American Meteorological Society*, 69(2):138–147.

Scherrer, S. C. and Appenzeller, C. (2014). Fog and low stratus over the Swiss Plateau-a climatological study. *International Journal of Climatology*, 34(3):678–686.

Schulz, H. M., Li, C.-F., Thies, B., Chang, S.-C., and Bendix, J. (2017). Mapping the montane cloud forest of Taiwan using 12 year MODIS-derived ground fog frequency data. *PLOS ONE*, 12(2):17.

Tambouratzis, T. and Tambouratzis, G. (2008). Meteorological data analysis using self-organizing maps. *International Journal of Intelligent Systems*, 23(6):735–759.

Troxler, F. and Wanner, H. (1991). Nebelkarten der Schweiz. *Geographica Helvetica*, 46(1):21–31.

Vautard, R., Yiou, P., and van Oldenborgh, G. J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nature Geoscience*, 2(2):115–119.

Veljović, K. and Vujović, D. (2018). Climatology of fog occurrence over a wide flat area in Serbia based on visibility observations. *International Journal of Climatology*, (September):1–14.

Wahl, S., Bollmeyer, C., Crewell, S., Figura, C., Friederichs, P., Hense, A., Keller, J. D., and Ohlwein, C. (2017). A novel convective-scale regional reanalysis COSMO-REA2: Improving the representation of precipitation. *Meteorologische Zeitschrift*, 26(4):345–361.

Wanner, H. (1979). Die Nebelverhältnisse im Querschnitt Jura-Alpen. *Mitteilungen der Naturforschenden Gesellschaft in Bern*, 36:57–62.

Willett, H. C. (1928). Fog and Haze, Their Causes, Distribution, and Forecasting. *Monthly Weather Review*, 56(11):435–468.

Chapter 6

Summary, Conclusion and Outlook

6.1 Summary and Conclusion

Fog has many economic as well as ecological impacts and it directly affects human life in many ways. The large number of fog influence factors shows that a comprehensive understanding of its causes and a precise mapping of the spatio-temporal distribution patterns are of great interest. Since there are justifiable concerns about the general applicability of existing fog retrieval methods, this thesis investigates new techniques of satellite based fog detection and the derivation of spatio-temporal information on fog distribution in Europe. The central novelties of this study are:

- No static assumptions about microphysical properties were used during fog retrieval.
- A novel hybrid approach based on machine learning methods was developed that can be continuously applied 24 hours a day.
- The algorithm covers all fog types. Areas of different fog types could also be differentiated indirectly from the generated product due to their typical diurnal and annual frequency cycles.
- For the first time, fog frequency maps for Europe could be produced for different general weather situations separately for each fog type.

In order to identify the shortcomings of existing physics-based fog retrieval methods and to check the validity of the theoretical assumptions used in these methods, an in situ measurement campaign was conducted at the Marburg ground truth and profiling station in Linden, Germany. Meteorological properties and droplet spectra were recorded over the complete vertical fog profile during several radiation fog events by means of a tethered balloon borne system and hypothesis **H1** was formulated as follows:

H1 Microphysical properties change during fog evolutionary stages. Static subadiabatic approximations as currently used for the differentiation between fog and lifted cloud layers are therefore not suitable to cover the complete fog life cycle.

The results of the measurement campaign were compared to theoretical assumptions used in current fog retrieval techniques in order to check for their applicability. An in-depth analysis showed that the theoretical assumption of a subadiabatic liquid water content profile in radiation fogs is valid in principle. However, strong variations during the different fog evolutionary stages indicate that the static assumptions of a subadiabatic liquid water content profile do not always apply. Hypothesis **H1** was

therefore confirmed. It is consequently necessary to adapt existing fog retrieval algorithms for a dynamic component that is able to account for changes in fog microphysics in order to gain a meaningful fog climatological data set (Chapter 2).

Thus, the main objective of this thesis was to develop a new fog retrieval technique based on satellite remote sensing data that would be able to cope with the complex meteorological and microphysical dynamics of fog and its spatio-temporal high variability in a satisfactory manner. Hypothesis **H2** was formulated accordingly:

H2 To overcome the problems of existing fog retrievals, a dynamic hybrid approach is needed that combines satellite and station data via statistical machine learning methods in order to derive fog information regardless of fog type or evolutionary stage and for the complete diurnal cycle.

The detection of low stratiform cloud layers in existing procedures as a preliminary stage towards fog retrieval is not yet globally applicable due to changes in the droplet spectrum during different fog life cycle stages, changes in land cover, in the solar position and in the satellite viewing angle. Thus, the first step was to dynamically extend existing low stratus detection approaches. In this context, the cloud detection was adapted so that the derivation of the thresholds, which are used to classify pixels into cloudy and non-cloudy, is done separately for day/night and land/sea. These thresholds are recalculated for each scene and a moving-window method additionally ensures that changes in the satellite viewing angle can be accounted for in the underlying histogram analysis. The small droplet proxy test which is used to differentiate low stratus from large droplet cloud types during daytime, is also extended by a moving window approach to ensure a fine grained threshold determination. This dynamical adaptation allows for a much more accurate time-independent area-specific low stratus recognition. The performance of the adapted algorithm was checked in an extensive validation procedure and the technique was deemed suitable for the derivation of a long term climatological data set (Chapter 3).

In order to circumvent the problems of existing fog retrieval approaches, which can only be applied during daytime and in radiation fog conditions but not during all fog life cycle stages and because of interference of multiple cloud layers with the satellite sensor, the next step was to implement a new fog detection algorithm using machine learning. For this purpose, a Random Forest approach was chosen. The algorithm was designed to predict cloud base altitudes for cloud covered regions and subsequently derive fog information via application of a statistically derived threshold separately for each pixel. The cloud base altitude information used for the training of the model was taken from METAR and SYNOP observations. By this means it was possible to develop an algorithm based on data of the Meteosat Second Generation satellite series that can give reliable information on the fog distribution in Europe for the complete diurnal cycle while not being limited to radiation fog types. The algorithm was validated using a leave-location-out cross validation procedure. The results show that the new algorithm can outperform existing procedures, especially when considering its broad applicability (Chapter 4). Hypothesis **H2** could therefore be confirmed as well.

After the successful development of a machine learning based fog retrieval technique and its application to a large data set between 2006 and 2015, hypothesis **H3** was tested:

H3 On the basis of a fog climatological data set derived from geostationary satellite data using a machine learning technique, it is possible to analyze the large-scale control mechanisms of fog dynamics spatially explicitly over the complete diurnal and annual cycle, with special consideration of the prevailing general weather conditions.

Because of its spatial coverage and its high temporal resolution, the computed fog product data set could be used to analyze fog distribution patterns in minute detail. In order to gain an overview of the spatio-temporal fog distribution in Europe, representative fog patterns were identified in a clustering procedure. Next, fog pattern occurrences were correlated to the occurrence of general weather situations in order to reveal the large-scale meteorological control mechanisms of fog dynamics. As the generated fog product provides spatially explicit fog distribution information every 15 minutes, it was possible to analyze the diurnal and annual fog frequency oscillation patterns in a last step. On the basis of the derived oscillation properties, regions with different fog types were successfully identified (Chapter 5). Hypothesis **H3** was thus also confirmed.

6.2 Outlook

After having summarized the results of this work, this section will give a brief outlook on future development plans and application potentials of the data and information gathered in this thesis.

The droplet spectrum data gained from the vertical profile measurements, which were carried out during the field measurement campaign, give valuable information on the dynamics of microphysical properties during different fog evolutionary stages. With the conduction of further profile measurements in different fog events it will be possible to derive more reliable conclusions about fog microphysics under changing conditions in different fog types. Because balloon-borne systems are, however, not operationally applicable, the data gathering could be realized using a 94/95 GHz FMCW radar (e.g. Delanoë et al., 2016; Huggard et al., 2008) as introduced in Chapter 2.2. These radar types are ideally suited for profiling fog layers as they are sensitive to small droplets that predominate in fog (Bennett et al., 2009; Maier et al., 2013). A feasibility study on the derivation of liquid water profiles using such a radar system was carried out by Thies et al. (2017) by investigating the influence of the drop size spectrum on the relationship between liquid water content and radar reflectivity. The results are promising and will be considered in future radar-based profile measurements of liquid water content in radiation fogs.

Additionally to the profile measurements, data sets on the distribution of low stratus clouds and fog were generated in the subsequent working packages. The information provided by these products is vital for the planning and expansion of transportation networks and photovoltaic installations. By applying the algorithm in nowcasting systems, real-time area-wide measures regulating traffic control could be put into practice. The products also give valuable information input for further investigations on fog dynamics and for studying the influence of different meteorological drivers of fog formation and dissipation.

As the fog retrieval algorithm developed in this context uses data that is available in quasi real-time, it is planned to develop an operational pipeline for the computation of the fog product in order to be able to continuously expand the analysis beyond the year 2015. With the help of an operational continuation, investigations of long-term trends in the frequency of fog occurrence, which were previously limited to station observation data (e.g. Avotniece et al., 2015; García-García and Zarraluqui, 2008; Srivastava et al., 2017; Troxler and Wanner, 1991; Vautard et al., 2009; Veljović and Vujović, 2018), can then be conducted spatially explicitly.

With the help of the high resolution visible band of the Meteosat Second Generation satellites, an inclusion of low earth orbit satellite systems like MODIS, a future adaptation of the algorithm to the soon to be deployed Meteosat Third Generation satellites and an increase in the station density, it will be possible to improve the spatial resolution of the product additionally. This will make it possible to capture even smaller-scale fog events and help determining the boundaries of detected fog layers more precisely. The planned incorporation of additional meteorological and geophysical data like air temperature and pressure fields from the regional COSMO-REA2 reanalysis product provided by the Hans-Ertel Center for Weather Research (Wahl et al., 2017) and land cover information available through the CORINE Land Cover project will also help to take into account smaller-scale meso- and micrometeorological effects in the product generation.

Bibliography

Avotniece, Z., Klavins, M., and Lizuma, L. (2015). Fog climatology in Latvia. *Theoretical and Applied Climatology*, 122(1-2):97–109.

Bennett, A., Nash, J., Gaffard, C., Moyna, B., Oldfield, M., and Huggard, P. (2009). Observations from the UK Met Office 94 GHz FMCW cloud radar. In *Proceedings of the 8th International Symposium on Tropospheric Profiling*, number October, pages 2–5, Delft.

Delanoë, J., Protat, A., Vinson, J.-P., Brett, W., Caudoux, C., Bertrand, F., Parent du Chatelet, J., Hallali, R., Barthes, L., Haeffelin, M., and Dupont, J.-C. (2016). BASTA: A 95-GHz FMCW Doppler Radar for Cloud and Fog Studies. *Journal of Atmospheric and Oceanic Technology*, 33(5):1023–1038.

García-García, F. and Zarraluqui, V. (2008). A fog climatology for Mexico. *Erde*, 139(1-2):45–60.

Huggard, P. G., Oldfield, M. L., Moyna, B. P., Ellison, B. N., Matheson, D. N., Bennett, A. J., Gaffard, C., Oakley, T., and Nash, J. (2008). 94 GHz FMCW cloud radar. *Proceedings of SPIE*, 7117:711704–711704–6.

Maier, F., Bendix, J., and Thies, B. (2013). Development and application of a method for the objective differentiation of fog life cycle phases. *Tellus, Series B: Chemical and Physical Meteorology*, 65(1):1–17.

Srivastava, S. K., Sharma, A. R., and Sachdeva, K. (2017). An observation-based climatology and forecasts of winter fog in Ghaziabad, India. *Weather*, 72(1):16–22.

Thies, B., Egli, S., and Bendix, J. (2017). The Influence of Drop Size Distributions on the Relationship between Liquid Water Content and Radar Reflectivity in Radiation Fogs. *Atmosphere*, 8(8):142.

Troxler, F. and Wanner, H. (1991). Nebelkarten der Schweiz. *Geographica Helvetica*, 46(1):21–31.

Vautard, R., Yiou, P., and van Oldenborgh, G. J. (2009). Decline of fog, mist and haze in Europe over the past 30 years. *Nature Geoscience*, 2(2):115–119.

Veljović, K. and Vujović, D. (2018). Climatology of fog occurrence over a wide flat area in Serbia based on visibility observations. *International Journal of Climatology*, (September):1–14.

Wahl, S., Bollmeyer, C., Crewell, S., Figura, C., Friederichs, P., Hense, A., Keller, J. D., and Ohlwein, C. (2017). A novel convective-scale regional reanalysis COSMO-REA2: Improving the representation of precipitation. *Meteorologische Zeitschrift*, 26(4):345–361.

Chapter 7

Zusammenfassung

Nebel hat eine große Anzahl wirtschaftlicher und ökologischer Auswirkungen und beeinflusst das menschliche Leben in vielerlei Hinsicht. Die Menge der Einflussfaktoren zeigt, dass ein umfassendes Verständnis seiner Bildungsursachen und eine genaue Erfassung seiner raumzeitlichen Verteilungsmuster von großem Interesse sind. Da es berechtigte Bedenken hinsichtlich der allgemeinen Anwendbarkeit bestehender Nebelerkennungsverfahren gibt, werden in dieser Arbeit neue Techniken der satellitengestützten Nebelerkennung untersucht und darauf aufbauend raumzeitliche Informationen über die Nebelverteilung in Europa abgeleitet. Die zentralen Neuentwicklungen dieser Arbeit sind:

- Bei der Nebelerkennung wurden keine statischen Annahmen bezüglich der mikrophysikalischen Eigenschaften herangezogen.
- Es wurde ein neuartiger hybrider Ansatz auf Basis von maschinellen Lernmethoden entwickelt, der 24 Stunden am Tag kontinuierlich angewendet werden kann.
- Der Algorithmus berücksichtigt alle Nebeltypen. Regionen mit unterschiedlichen Nebeltypen konnten auf Basis des erzeugten Nebelprodukts aufgrund ihrer typischen Tages- und Jahresgänge indirekt voneinander unterschieden werden.
- Erstmals konnten für Europa Nebelfrequenzkarten für verschiedene Großwetterlagen separat für jeden Nebeltyp erstellt werden.

Um die Unsicherheiten bestehender physikalischer Nebelerkennungsverfahren zu identifizieren und um die Gültigkeit der in diesen Verfahren verwendeten theoretischen Annahmen zu überprüfen, wurde eine Messkampagne an der Umweltbeobachtungs- und Klimafolgenforschungsstation in Linden durchgeführt. Meteorologische Eigenschaften und Tröpfchenspektren wurden über das gesamte vertikale Nebelprofil während mehrerer Strahlungsnebelereignisse mit Hilfe eines ballongestützten Systems aufgezeichnet. Hypothese **H1** wurde entsprechend wie folgt formuliert:

H1 Mikrophysikalische Eigenschaften ändern sich während der unterschiedlichen Entwicklungsphasen von Nebel. Statische subadiabatische Approximationen, wie sie derzeit zur Unterscheidung zwischen Nebel und gehobenen Wolkenschichten verwendet werden, sind daher nicht in der Lage, den gesamten Nebellebenszyklus zu repräsentieren.

Die Ergebnisse der Messkampagne wurden mit theoretischen Annahmen verglichen, die in aktuellen Nebelerkennungsverfahren verwendet werden. Eine detaillierte Analyse dieser Gegenüberstellung hat gezeigt, dass die theoretische Annahme

eines subadiabatischen Flüssigwassergehaltsprofils in Strahlungsnebeln zwar grundsätzlich gültig ist. Starke Schwankungen während der verschiedenen Entwicklungsstadien des Nebels deuten jedoch darauf hin, dass die statischen Annahmen des subadiabatischen Flüssigwassergehaltsprofils nicht immer zutreffen. Hypothese **H1** wurde daher bestätigt. Es ist folglich notwendig, bestehende Nebelerkennungsverfahren um eine dynamische Komponente zu erweitern, die in der Lage ist, Änderungen in der Nebelmikrophysik zu berücksichtigen, um einen sinnvollen nebelklimatologischen Datensatz erzeugen zu können (Kapitel 2).

Das Hauptziel dieser Arbeit ist entsprechend die Entwicklung eines neuen Nebelerkennungsverfahrens auf Basis von Satelliten-Fernerkundungsdaten, mit dessen Hilfe es möglich ist, die komplexe meteorologische und mikrophysikalische Dynamik von Nebel und seine raumzeitlich hohe Variabilität zu erfassen. Hypothese **H2** wurde entsprechend formuliert:

H2 Um die Probleme bestehender Nebelerkennungsverfahren zu überwinden, bedarf es eines dynamischen Hybrid-Ansatzes, der Satelliten- und Stationsdaten über statistische maschinelle Lernverfahren kombiniert, um daraus Nebelinformationen abzuleiten und zwar unabhängig vom Nebeltyp und der Evolutionsstufe, in der sich der Nebel befindet, sowie über den gesamten Tagesgang.

Da die Detektion von tiefliegenden Stratuswolken in bestehenden Verfahren als Vorstufe zur Nebelerkennung aufgrund von Veränderungen im Tröpfchenspektrum während verschiedener Nebellebensphasen, Änderungen der Landbedeckung, der Sonnenposition und des Satellitenblickwinkels noch nicht global einsetzbar ist, bestand der erste Schritt darin, existierende Ansätze zur Erkennung von tiefliegenden Stratuswolken dynamisch zu erweitern. In diesem Zusammenhang wurde die Wolkenerkennung so angepasst, dass die Ableitung der Schwellwerte, die zur Klassifizierung von Pixeln in wolkige und nicht-wolkige Bereiche dienen, getrennt für Tag/Nacht und Land/See erfolgt. Diese Schwellwerte werden für jede Szene neu berechnet, und ein “moving window“ Verfahren stellt zusätzlich sicher, dass Änderungen des Satellitenblickwinkels in der zugrunde liegenden Histogrammanalyse berücksichtigt werden können.

Ein weiterer Prozessierungsschritt, mit dessen Hilfe tagsüber niedrige Stratuswolken von höheren Wolken mit großen Tropfen unterschieden werden können, wurde ebenfalls durch einen “moving window“ Ansatz erweitert, um eine genauere Schwellwertbestimmung zu gewährleisten. Diese dynamische Anpassung ermöglicht eine deutlich präzisere, zeitunabhängige und regionsspezifische Erkennung von niedrigem Stratus. Die Qualität des so erweiterten Algorithmus wurde in einem umfangreichen Validierungsverfahren überprüft. Die Ergebnisse dieser Validierung haben gezeigt, dass das entwickelte Verfahren für die Ableitung eines langfristigen klimatologischen Datensatzes geeignet ist (Kapitel 3).

Um die übrigen Probleme der bestehenden Nebelerkennungsverfahren zu umgehen, die nur tagsüber bei Strahlungsnebelbedingungen und nur bei bestimmten Phasen des Nebellebenszyklus anwendbar sind, wurde im nächsten Schritt ein neuer Nebelerkennungsalgorithmus auf Basis von maschinellen Lernverfahren implementiert. Zu diesem Zweck wurde ein sog. “Random Forest“ Ansatz gewählt. Der Algorithmus wurde so entwickelt, dass dieser die Höhen der Wolkenunterseiten für wolkenbedeckte Regionen vorhersagt, damit anschließend Informationen zur Nebelverteilung durch Anwendung eines statistisch abgeleiteten Schwellwerts für jedes Pixel separat abgeleitet werden können. Die für das Training des Modells verwendeten Daten zur Wolkenhöhe stammen aus METAR- und SYNOP-Beobachtungen. Auf diese Weise ist es gelungen, einen

Algorithmus zu entwickeln, der basierend auf Daten der MSG-Satellitenreihe zuverlässige Informationen über die Nebelverteilung in Europa für den gesamten Tagesgang und alle Nebeltypen liefert. Die Qualität des Algorithmus wurde mit Hilfe eines Kreuzvalidierungsverfahrens geprüft. Die Ergebnisse zeigen, dass der neue Algorithmus bestehende Verfahren übertrifft, insbesondere wenn man seine breite Anwendbarkeit berücksichtigt (Kapitel 4). Hypothese **H2** konnte daher auch bestätigt werden.

Nach der erfolgreichen Entwicklung eines Nebelerkennungsalgorithmus auf Basis von maschinellen Lernverfahren und dessen Anwendung auf einen umfangreichen Datensatz zwischen 2006 und 2015 wurde Hypothese **H3** getestet:

H3 Auf der Grundlage eines nebelklimatologischen Datensatzes, der mit Hilfe maschineller Lernverfahren aus geostationären Satellitendaten abgeleitet wurde, ist es möglich, die großräumigen Kontrollmechanismen der Nebeldynamik unter besonderer Berücksichtigung der vorherrschenden allgemeinen Wetterbedingungen flächendeckend über den gesamten Tages- und Jahresgang zu analysieren.

Aufgrund seiner räumlichen Abdeckung und seiner hohen zeitlichen Auflösung konnte das erstellte Nebelprodukt dazu verwendet werden, Nebelverteilungsmuster im Detail zu analysieren. Um einen Überblick über die raumzeitliche Nebelverteilung in Europa zu erhalten, wurden in einem Klassifizierungsverfahren repräsentative Nebelverteilungsmuster identifiziert. Anschließend wurde das Auftreten dieser Nebelverteilungsmuster mit dem Auftreten bestimmter Großwetterlagen korreliert, um die großräumigen meteorologischen Kontrollmechanismen der Nebeldynamik zu erfassen. Da das erzeugte Nebelprodukt alle 15 Minuten flächendeckende Informationen zur Nebelverteilung liefert, war es möglich, in einem letzten Schritt die täglichen und jährlichen Schwankungen in der Nebelfrequenz zu analysieren. Auf Basis der ermittelten Schwankungsmuster konnten anschließend Regionen mit unterschiedlichen Nebeltypen identifiziert werden (Kapitel 5). Hypothese **H3** wurde damit ebenfalls bestätigt.

Mit Hilfe der in dieser Arbeit erzeugten Datensätze konnten wichtige Informationen zur Dynamik der mikrophysikalischen Eigenschaften von Nebel zur Verfügung gestellt werden. Die flächenhaften Produktdatensätze zu niedriger Stratusbewölkung und Nebel können unter anderem bei der Planung und dem Ausbau von Verkehrsnetzen und Photovoltaikanlagen Berücksichtigung finden. Durch den Einsatz des Nebelerkennungsverfahrens in Nowcasting-Systemen können flächendeckende Maßnahmen zur Verkehrsregulierung mit kurzer Verzögerung umgesetzt werden. Die Produkte liefern außerdem wertvolle Informationen für weitere Untersuchungen zur Nebeldynamik und zur Untersuchung des Einflusses verschiedener meteorologischer Treiber die bei der Nebelbildung und -auflösung eine Rolle spielen.

Mit Hilfe des hochaufgelösten HRV-Kanals der MSG-Satelliten, durch das Miteinbeziehen von erdnahen Polarorbitern wie MODIS, durch die zukünftige Anpassung des Algorithmus an die bald zum Einsatz kommende dritte Generation der Meteosat-Reihe und durch eine Erhöhung der Stationsdichte wird man in der Lage sein, die räumliche Auflösung des Produktes weiter zu verbessern. Damit wird es möglich sein, auch kleinskaligere Nebelereignisse zu erfassen und Grenzbereiche von Nebelschichten genauer bestimmen zu können. Außerdem soll die geplante Einbeziehung zusätzlicher meteorologischer und geophysikalischer Daten wie etwa Lufttemperatur- und Druckfelder aus regionalen Reanalyseprodukten und Landbedeckungsinformationen aus dem CORINE-Projekt zukünftig dazu beitragen, kleinere meso- und mikrometeorologische Effekte in der Produktgenerierung berücksichtigen zu können.

Erklärung

Ich versichere, dass ich meine Dissertation

Satellite-Based Fog Detection:
A Dynamic Retrieval Method for Europe Based on Machine Learning

selbständig, ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat auch noch keinen sonstigen Prüfungszwecken gedient.

Marburg a. d. Lahn, den 10. Januar 2019

Sebastian Egli